

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 26, 2003, 11:37:36 ; Search time 1160 Seconds  
(without alignments)  
2918.186 Million cell updates/sec

Title: US-09-600-714-41

Perfect score: 1254

Sequence: 1 atgagcttaagctaccgcg.....attgcctgttgatttaa 1254

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1254	100.0	1254	20	AAx86522	cDNA sequence of t
2	177	14.1	1805	24	ABK12333	cDNA encoding huma
3	175.4	14.0	1792	24	ABZ11480	Human polynucleoti
4	165.4	13.2	1853	21	AAH62486	Human membrane-ass
5	165.4	13.2	1949	22	AAH25763	Oesophagus cancer
6	165.4	13.2	1952	24	ABK49215	cDNA encoding huma
7	165.4	13.2	1968	25	ABX76257	Lung cancer-associ
8	165.4	13.2	1970	21	AAH76344	Human ORF1899

c	9	159.4	12.7	1679	24	ABK12334	cDNA encoding mous
c	10	152.4	12.2	3706	22	AAH18728	Human cDNA sequenc
c	11	150.2	12.0	2098	24	ABK49216	cDNA encoding mous
c	12	139.8	11.1	486	22	ABA42886	Human breast cell
c	13	139.8	11.1	486	22	ABA53307	Human foetal liver
c	14	139.8	11.1	486	22	ABA23081	Probe #1547 for ge
c	15	139.8	11.1	486	22	AAK01568	Human brain expres
c	16	139.8	11.1	486	22	AAK27012	Human bone marrow
c	17	139.8	11.1	486	22	AAI11615	Probe #1548 for ge
c	18	139.8	11.1	486	22	AAI32911	Probe #1597 used t
c	19	139.8	11.1	486	22	AAI01542	Probe #1533 used t
c	20	139.8	11.1	486	23	ABS26596	Human liver single
c	21	139.8	11.1	486	24	ABS01595	Human genome-deriv
c	22	120	9.6	1571	24	ABL90445	Human polynucleoti
c	23	113.8	9.1	1840	23	ABA48015	Drosophila melanog
c	24	111.8	8.9	123	22	ABA65900	Human breast cell
c	25	111.8	8.9	123	22	ABA32986	Human foetal liver
c	26	111.8	8.9	123	22	AAK14315	Probe #11452 for g
c	27	111.8	8.9	123	22	AAK40045	Human brain expres
c	28	111.8	8.9	123	22	AAI20829	Human bone marrow
c	29	111.8	8.9	123	22	AAI46064	Probe #14750 used
c	30	111.8	8.9	123	22	AAI06540	Probe #6531 used t
c	31	111.8	8.9	123	22	ABS39627	Human liver single
c	32	111.8	8.9	123	23	ABS14127	Human genome-deriv
c	33	111.8	8.9	123	24	AAZ97083	Human secreted pro
c	34	109.6	8.7	1650	21	AAI90968	Human polynucleoti
c	35	108.6	8.7	484	22	ABL33617	Human immune syste
c	36	108.6	8.7	16918	24	AAH28951	Human MOLT CDNA.
c	37	101.2	8.1	1765	24	AAH98147	Human EST-derived.
c	38	93.2	7.4	1222	22	AAH3616	Human immune syste
c	39	82.2	6.6	16918	24	AAH31425	Human secreted pro
c	40	79.2	6.3	730	22	ABL90632	Human polynucleoti
c	41	79.2	6.3	730	24	ABN75889	Human glycoprotein
c	42	61.2	4.9	198	24	AAH93284	cDNA encoding SRT
c	43	57	4.5	449	22	ABA50433	Human breast cell
c	44	56.6	4.5	446	22	ABA68382	Human foetal liver
c	45	56.6	4.5	446	22		

#### ALIGNMENTS

RESULT 1	AAx86522	standard: cDNA: 1254 BP.
ID	AAx86522	
AC	AAx86522	
XX		
DT	04-OCT-1999	(first entry)
XX		
DE	cDNA sequence of the prevalent allele of the Rhd gene.	
XX		
KW	Allele: Rhesus D antigen: RHD; weak D phenotype; blood transfusion; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1254
FT		/tag= a
XX		
PN	W09937763-A2.	
XX		
PD	29-JUL-1999.	
XX		
PF	18-DEC-1998:	98WO-EP08319.
XX		
PR	23-JAN-1998:	98BP-0101203.
XX		
PA	(DRKB-) DRK BLUTSPENEDIENST BADEN WUERTTEMBERG.	
XX		
PI	Flegel WA, Wagner FF;	
XX		
DR	WPI: 1999-469127/39.	

DR P-PSDB: AAY24056.  
 XX Nucleic acid sequences correlated with Rhesus weak D phenotype.  
 PT useful for screening blood from donors and recipients for  
 PT transfusion methods  
 XX  
 PS Disclosure: Fig 2; 64pp; English.  
 XX  
 CC The present sequence represents the prevalent allele of the Rhesus D  
 CC (Rhd) antigen gene. The specification describes a Rhd contributing to  
 CC or indicative of the weak D phenotype, where the Rhd polypeptide  
 CC carries at least one missense mutation as compared to the wild-type Rhd,  
 CC in its transmembrane and/or intracellular regions, especially in amino  
 CC acid positions 2-16, 114-149, 179-225 or/and 267-397, with the proviso  
 CC that the D antigen does not carry a single missense mutation leading to  
 CC a F223V or F283I substitution. The probes and antibodies are useful in  
 CC the methods for detection of weak D phenotypes. Red blood cells, from  
 CC probands, are useful for the assessment of the affinity, avidity and/or  
 CC reactivity of monoclonal anti-D antibodies, polyclonal anti-D antisera or  
 CC of anti-globulin or anti-human-globulin antisera. Detecting the presence  
 CC of the Rhd associated with weak D phenotype is useful for determining  
 CC that a patient in need of a blood transfusion is to be transfused with  
 CC RhD negative blood from a donor. Alternatively, testing for weak D  
 CC phenotype Rhd in the blood of a donor is useful for determining whether  
 CC the donor blood should be excluded for transfusion to patients having  
 CC wild type Rhd or weak D types, other than that of the donor weak D  
 CC type.  
 CC  
 XX  
 SQ Sequence 1254 BP; 241 A; 322 C; 351 G; 340 T; 0 other;  
 Query Match 100.0%; Score 1254; DB 20; Length 1254;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 601 ACAGCAAGCATACCAAGTTTGTCTGCCATGCTGGGCGCCCTCTTCTGTGATGTTCTGG 660  
 DB 601 ACAGCAAGCATACCAAGTTTGTCTGCCATGCTGGGCGCCCTCTTCTGTGATGTTCTGG 660  
 QY 661 CCAAGTTTCAACTCTGCTGTGAGAAAGTCCAAATGGAAGAATGCCGTGTTCAAC 720  
 DB 661 CCAAGTTTCAACTCTGCTGTGAGAAAGTCCAAATGGAAGAATGCCGTGTTCAAC 720  
 QY 721 ACCTACTATGCTGTAGACAGTACGCTGGTGGACAGCCATCTCAGAGTATCCTTGGGTCAC 780  
 DB 721 ACCTACTATGCTGTAGACAGTACGCTGGTGGACAGCCATCTCAGAGTATCCTTGGGTCAC 780  
 QY 781 CCCCAAGGAAGATCAGCAAGACTTATGTGCACAGTCCGCTGTGTGGCAGAGGCGTGGCT 840  
 DB 781 CCCCAAGGAAGATCAGCAAGACTTATGTGCACAGTCCGCTGTGTGGCAGAGGCGTGGCT 840  
 QY 841 GTGGGTAACCTCGTGTACCTGATCCCTTTCCTGGCTGGCTATGCTGCTGGTCTTGTG 900  
 DB 841 GTGGGTAACCTCGTGTACCTGATCCCTTTCCTGGCTGGCTATGCTGCTGGTCTTGTG 900  
 QY 901 GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTCCGGGGTGTGTAACGAGTGTG 960  
 DB 901 GCTGGGCTGATCTCCGTCGGGGAGCCAAAGTACCTCCGGGGTGTGTAACGAGTGTG 960  
 QY 961 GGGATTCCCCACAGCTTCATATGGCTACACTTATGCTGCTGGCTCTGTGGAGAG 1020  
 DB 961 GGGATTCCCCACAGCTTCATATGGCTACACTTATGCTGCTGGCTCTGTGGAGAG 1020  
 QY 1021 ATCATCTACATTTGTGGTGGCTGGTGGTGTGATACGCTGGAGCCGGAATGGCATGATTGGC 1080  
 DB 1021 ATCATCTACATTTGTGGTGGCTGGTGGTGTGATACGCTGGAGCCGGAATGGCATGATTGGC 1080  
 QY 1081 TTCCAGGTCCTCCTCAGCATTTGGGGAAGTACGCTTGGCATCTGTATAGCTCTCAGCTGT 1140  
 DB 1081 TTCCAGGTCCTCCTCAGCATTTGGGGAAGTACGCTTGGCATCTGTATAGCTCTCAGCTGT 1140  
 QY 1141 GGTCTCCTGACAGGTTTGTCTCTAAATCTTAAATATGGAAGACACCTCATGAGGCTTAA 1200  
 DB 1141 GGTCTCCTGACAGGTTTGTCTCTAAATCTTAAATATGGAAGACACCTCATGAGGCTTAA 1200  
 QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTGGATTTTAA 1254  
 DB 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCCATTTGGCTGTGGATTTTAA 1254

RESULT 2  
 ABK12333  
 ID ABK12333 standard; cDNA; 1805 BP.  
 XX  
 AC ABK12333;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE cDNA encoding human nonerythroid Rh glycoprotein RhBG.  
 XX  
 KW Human: nonerythroid Rh glycoprotein; Rhbg; RhBG; antibody; immunogen;  
 KW Rh type B transporter gene; polytypic transporter -type protein;  
 KW ion transporter; chromosome 1q21.3; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 39..1415  
 FT /tag= a  
 FT /product= "Human RhBG protein"  
 FT misc\_feature 1248..1377  
 FT /tag= b  
 FT /note= "Encodes C'-tail"  
 FT polyA\_signal 1769..1774  
 FT /tag= c  
 FT /standard\_name= "PolyA signal"  
 FT /note= "Atypical polyadenylation site"







Best Local Similarity 49.1%; Pred. No. 2.8e-37;  
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

```
QY 141 CTATCAAGTTGGCCAGATCTGACGTGATGGCGGCGCTTGGGCTTCTCCACCTC 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 CTACCCAGCTTCAGAGAGCTGACAGTGTGCTTCGTGGGCTTCTCCATGAC 276
QY 201 GAGTTTCCGGAGACAGCTGAGAGTGTGGCGTTCAACCTTTCATGCTGGCGCTTG 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 TTTCCTGACAGCCCTTACGGCTTCAGCGCCCTGGGCTTCAACTTCCCTGTGACAC 336
QY 261 TGTGACGTGGGCAATCTGCTGACCGCTTCCGACACAGTTCCTTCTGGAGAGTGT 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 CATCCAGTGGGCGCTCTCATGACGGCTGGTTTCACTTCTTAAACAAGCCGCTACAT 396
QY 321 CATCACACTGTTCAAGATTGCGCTGGCCACATGATGCTTTCGGCTGATCTCAGT 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 397 CGTGGGCGGTGGAGAACCTCATCAAGCTGATCTTCGCTGGCTCTGTCTGGCGCTT 456
QY 381 GGATGCTGTCTGGGAGAGTCAACTTGGCGCAGTTGGTGGTATGCTGTGGTGGAGT 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 457 TGGGGCAGTTCTGGTAAAGTACGCCCATTCAGTCTGATCATCATCTTCTTCCAGT 516
QY 441 GACAGCTTAGGCAACCTGAGATGCTCATCACTATATCTTCAACACAGACTACACAT 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 517 GACCTCTTCTGCTGTAATGATGATCTCTCTTAACTGCTTAAAGTGAAGGATGAC 576
QY 501 GAACATGATGACATCTACGCTGTTCCAGCTATTTTGGGCTGTGGCGCTGGCT 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577 AGGCTCATGATGACATCCACATTTTGGCGCTTACTTGGGCTCAGAGTACCGGATCT 636
QY 561 GCCAAGCCTCTACCCGAGGAGAACGAGATTAAGATCAGACAGACATACCCAGTT 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 637 CTACCGACGCAACTGAGAGAGAGCAAGAGAGAAATTCCTGTACAGTGGAGCT 696
QY 621 GTCTGCATGCTGGGGCCCTCTCTGTGATGATGTTCTGGCCAGTTTCAACTCTCT 680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 697 CTTGGCATGATGGACCCCTCTCTGTGAGATGATGGCCAGCTTCAACTCAGCCT 756
QY 681 GCTGAGAGTCCAAATCGAAGAGAAAGATCCGTTTCAACACCTTATGCTAGTAC 740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 757 ATCTACCAATGGGGAACAGCAGCAGCAGCGCATCAACACTACTGCTCTGGCAGC 816
QY 741 CAGCTGTGTGACAGCCATCTCAGGTCATCTTGGCTCACCCCAAGAGATCAGCAA 800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 817 CTCCTGCTTACCTCGTGGCATATTCAGTCCCTCAGCAAGAAAGGCAAGCTGGACAT 876
QY 801 GACTTATGTGACAGTGGCGGTGTGGCAGAGAGCGTGGGATACCTCGTGCACCT 860
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 877 GGTGCAATCCAGAAATGCCACGCTCCAGAGAGGGGTGGCTGTACCGCTGTAGAT 936
QY 861 GATCCCTTCTCGGTGGCTTGCATGTGCTTGGTCTTGTGGCTGGGCTGATCTCCGTC 920
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 937 GATGCTCATGCTTACGTTGCCCTCATTCATTCGCGCTTGTGGGCTCATCTCCACCT 996
QY 921 GGGAGCCAGTACTCCGGGGGTGTGTGAACGAGTGTGGGATTTCCCAACAGCTCAT 980
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 997 GGGTTTGTATATCTGACCCCATTCCTGAGATGCCGTCGACATCCAGACATCATGTG 1056
QY 981 CATGGCTACACTTCACTGCTGTGGTGTGCTTGGAGAGATCATATCATATGTG 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1057 CATTAACAATCTGATGCGATTCCTGTGGCATCATAGCGGCGCATGTGGGTGTG 1111
```

RESULT 6  
ABK49215  
ID ABK49215 standard; cDNA; 1952 BP.

AC ABK49215;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE cDNA encoding human Rh type C gene (RHCG) protein.

```
XX RHCG: human; non-erythroid Rh type C glycoprotein;  
KW chromosome 15q25; gene; ss.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH CDS 24..1440  
FT /*tag= a  
FT /product= "RHCG protein"  
FT polyA_signal 1885..1890  
FT /*tag= b  
PN WO200220719-A2.  
XX  
PD 14-MAR-2002.  
XX  
PF 05-SEP-2001: 2001MO-US27503.  
XX  
PR 07-SEP-2000: 2000US-230660P.  
XX  
PA (NYBL-) NEW YORK BLOOD CENT INC.  
XX  
PI Huang C, Liu Z;  
XX  
XX WPI: 2002-351774/38.  
DR P-PSDB; AAU78997.  
XX  
PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C  
PT and glycoproteins which have a characteristic twelve transmembrane  
PT domain structure -  
XX  
XX Claim 2: Fig 1; 53pp; English.  
XX  
CC This invention relates to the nucleic acid and protein sequences of  
CC novel human and mouse non-erythroid Rh type C glycoprotein (RhCG).  
CC The RhCG protein and the mouse homologue (rhcg) have a characteristic  
CC 12 transmembrane domain structure and are expressed in kidneys and  
CC testis. The invention also comprises a method for antibody that  
CC specifically binds an epitope of the glycoprotein and a method for  
CC detecting the protein using this antibody. The antibodies of the  
CC invention may be used in Western blots, enzyme linked immunosorbent  
CC assays (ELISAs) or immunohistochemical assays to identify the non-  
CC erythroid tissues, particularly kidney and testis, that express the  
CC RhCG or Rhcg glycoproteins. The methods are used for detecting an  
CC Rhcg or and Rhcg glycoprotein in a sample. The present sequence  
CC represents the cDNA encoding the human RHCG Rh type C glycoprotein  
CC (Rhcg) protein sequence of the invention. The gene encoding this  
CC protein is located on human chromosome 15q25.  
XX  
SQ Sequence 1952 BP; 398 A; 586 C; 524 G; 444 T; 0 other;  
Query Match 13.2%; Score 165.4; DB 24; Length 1952;  
Best Local Similarity 49.1%; Pred. No. 2.8e-37;  
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;
```

```
QY 141 CTATCAAGTTGGCCAGATCTGACCGTATGGCGGCGCTTGGGCTTCTCCACCTC 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 CTACCCAGCTTCAGAGAGCTGACAGTGTGCTTCGTGGGCTTCTCCATGAC 260
QY 201 GAGTTTCCGGAGACAGCTGAGAGTGTGGCGTTCAACCTTTCATGCTGGCGCTTG 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 TTTCCTGACAGCCCTTACGGCTTACGGCGCTTGGGCTTCAACTTCCCTGTGACAC 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 261 TGTGAGTGGGCAATCTGCTGAGCGCTTCTGAGACCGATTCCTTCTGGAGAGTGT 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 CATCCAGTGGGCGCTGCTCATCAAGGCTTGTCCACTTCTTAAACAAGCCGCTACATGCT 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 321 CATCAACTGTTCAAGTATTCGCTGGCCACCATGATGCTTGTGGTGGTCAATCTCACT 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 381 CGTGGGCGGTGGAGAACCTCATCAAGCTGATCTTGTGCGTGGCTGTCTGTGGCGCTT 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 381 GGATGCTGTCTTGGGAGAGGTCAACTTGGCGCAGTTGGTGTGATGCTGTGGAGAGT 440
```

Db 441 TGGGGCAGTCTGGGTAAGTCAGCCCATTCAGCTGCATCATGACTTCTTCCAGT 500  
 QY 441 GACAGCTTAGGCAACCGAGAGATGTCATCATATATCTTCAACACAGACTATTACACAT 500  
 Db 501 GACCCCTTCCTGCTGAATGAGTTTCTTCTTAACCTGCTAAAGGTAAGGATGACAG 560  
 QY 501 GAACATGATGACATCTAGCTGCTGAGCCATATTTGGGCTGTCTGGCTGTGCT 560  
 Db 561 AGGCTCCATGACCATTCACACATTTGGGCTTCTTGGGCTCACAGTACCCGATCCT 620  
 QY 561 GCCAAAGCCTTACCCGAGGAGACGAGAGATTAAGATGACAGACGATACCGATT 620  
 Db 621 CTACGACGCAACCTAGACAGAGACAGAGAGAGATCTGTACCACTGCGACCT 680  
 QY 621 GTCGCAATGCTGGGCGCCCTTCTTGTGATGTTTGGCCCACTTCACTGCTCT 680  
 Db 681 CTTCGCAATGATGGACCCCTTCTTCTGTGATGATGAGCCCTTCACTCACTGACCAT 740  
 QY 681 GCTGAGAGTCCATCGAAAGAGAGATGCGCTGTTCAACACTTACTATGCTGTAGCAGT 740  
 Db 741 ATCCATGATGAGGAGACACGACGACGACGACGACGACGACGACGACGACGACG 800  
 QY 741 CAGCGTGTGACAGGCAATCTCAGGGTCAATCTTGGCTCAACCCCAAGGAGATGACGAA 800  
 Db 801 CTGCGTGTGACGCTGAGGCAATCTCAGGGTCAATCTTGGCTCAACCCCAAGGAGATGACGAA 800  
 QY 801 GACTTATGTCACATCTGCGGTGTGTGGCAGAGAGGCTGCTGCTGCTGCTGCTGCT 860  
 Db 861 GGTGACATTCAGAAATGCGACCCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 920  
 QY 861 GATCCCTTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920  
 Db 921 GATGCTCATGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980  
 QY 921 GGGAGCCAGTACCTGCGGGGTGTGTACCCAGAGTCTGCTGCTGCTGCTGCTGCTGCT 980  
 Db 981 GGGTGTGCTATACCTGAGCCCATTCCTGAGTCCCGGCGACATTCAGAGACATGATG 1040  
 QY 981 CATGGGCTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040  
 Db 1041 CATTAACAATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1095

RESULT 7  
 ABX76257 standard; DNA; 1968 BP.

AC ABX76257;  
 XX  
 DE 02-APR-2003 (first entry)  
 XX  
 Lung cancer-associated polynucleotide #123.  
 XX  
 Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;  
 KW antiinflammatory; antiaesthetic; non-small cell lung cancer; atelectasis;  
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
 XX  
 OS Unidentified.  
 PN WO200286443-A2.  
 PD 31-OCT-2002.  
 XX  
 18-APR-2002; 2002WO-US12476.  
 PF 18-APR-2001; 2001US-284770P.  
 PR 10-MAY-2001; 2001US-290492P.  
 PR 09-NOV-2001; 2001US-339245P.  
 PR 13-NOV-2001; 2001US-350666P.  
 PR 29-NOV-2001; 2001US-334370P.

PR 12-APR-2002; 2002US-372246P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Aziz N, Murray R;  
 XX  
 DR WPI: 2003-093161/08.  
 DR P-PSDB; AB056530.  
 PT Detecting a lung cancer-associated transcript in a cell from a patient  
 PT for treating lung cancer, by contacting a biological sample from the  
 PT patient with a polynucleotide that exhibits increased or decreased  
 PT expression in lung cancer  
 PS Claim 22; Page 285-286; 453pp; English.

CC The invention relates to a method for detecting a lung cancer-associated  
 CC transcript in a cell from a patient, comprising contacting a biological  
 CC sample from the patient with a polynucleotide that selectively hybridises  
 CC to a sequence that is at least 80 % identical to a gene that exhibits  
 CC increased or decreased expression in lung cancer samples. Lung  
 CC cancer-associated polynucleotides and polypeptides are used for  
 CC identifying a compound that modulates a lung cancer-associated  
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated  
 CC cell to treat lung cancer in a patient and for treating a mammal having  
 CC lung cancer by administering a modulatory compound identified. The  
 CC methods are useful for treating lung cancer, such as small cell lung  
 CC cancer, non-small cell lung cancer or other benign or precancerous  
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive  
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial  
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides  
 CC and polypeptides are useful for diagnostic purposes and as targets for  
 CC screening for therapeutic compounds that modulate lung cancer, such as  
 CC antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated  
 CC polynucleotides of the invention.

SQ Sequence 1968 BP; 400 A; 592 C; 529 G; 445 T; 2 other;

Query Match 13.2%; Score 165.4; DB 25; Length 1968;  
 Best Local Similarity 49.1%; Pred. No. 2.8e-37;  
 Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCAAGATCTGACGCTGATGAGGCGCATTTGGGCTTCCCTACCTC 200  
 Db 217 CTACCAAGCTTCCAGAGCTGACGCTGATGATGCTTGGGCTTCCCTACATGAC 276  
 QY 201 GAGTTCCGAGACACAGCTGAGAGAGTGGCTTCAACCTTCTATGCTGGCGCTTGG 260  
 Db 277 TTTCTGCGAGCGCTTACGGCTTACGCGCGCTTCAACTTCTGTTGGACGCTTGG 336  
 QY 261 TGTGCAAGTGGCAATCTGCTGAGCGCTTCTGAGCCAGTTCCTTGGAGAGTGGT 320  
 Db 337 CATCCAGTGGGCGCTCTCATGAGGCGTGGTTCATCTTCAAGACCGCTACATCGT 396  
 QY 321 CATCAACAGCTTCAAGATTCGCTGCGCCACCATGAGTGGCTTGGCTGATCTCAGT 380  
 Db 397 CGTGGGCGTGGAGAACCTTATCAAGCTGATCTTCTGCGGCTTCTGCTGGGCTT 456  
 QY 381 GGAATGCTTCTTGGGAGAGTCAACTTGGCGCAGTTGGTGGATGAGTGGTGAAGT 440  
 Db 457 TGGGGCAGTTCTGGGTAAGTAGGCCCATTCAGCTGCTCATGACTTCTTCCAACT 516  
 QY 441 GACAGCTTATGAGCACTGAGATGCTCATATATCTTCAACACAGACTTACACAT 500  
 Db 517 GACCCCTTCCTGCTGATGATGATTCATCTTCAACCTGTAAGGTAAGGATGACAG 576  
 QY 501 GAACATGATGACATCTAGCTGCTGAGCCATATTTGGGCTGTCTGGGCTGTGCT 560  
 Db 577 AGGCTCCATGACCATTCACACATTTGGGCTTCTTGGGCTCACAGTACCCGATCCT 636  
 QY 561 GCCAAAGCCTTACCCGAGGAGACGAGAGATTAAGATGACAGACGATACCGAGTTT 620  
 Db 637 CTACGACGCAACCTAGACGAGGCAAGAGAGACAGAAATTTCTGTACACGTCGACCT 696







Db 759 ATCTACATGGGAGACAGCAGCAGCCGCGCATCAACCTACTGCTCTTGCGAGC 818  
QY 741 CAGCGTGGTACAGCCATCTCAGGGTCATCTTGCTGACCCCAAGGAGATCAGCAA 800  
Db 819 CTGGTGCTTACCTCGGTGGCAATATCCAGTCCCTGCACAGAAGGACAGCTGGACAT 878  
QY 801 GACTTATGTGCACAGTCCGGTGTGGAGAGAGGGGTGGCTGTGACTGCTGCACCT 860  
Db 879 GGTCACATCCAGATGACAGCTCGCAGAGGGGTGGCCCTGGGTACCGCTGCAGAT 938  
QY 861 GATCCCTTCTCCGCTGCTGCCATGCTGCTGGTCTGCTGGCTGATCTCCGTCGG 920  
Db 939 GATGCTCATGCTTACGGTGGCCATCATCGCTGCTGCTGGCATCATCTCACCCCT 998  
QY 921 GGGAGCCAGTACCTGGCGGGGTGTGTAAACGAGTGTGGGATTCGCCACAGCTCAT 980  
Db 999 GGGTTTGTATACCTGACCCCATCTCTGAGTCCCGGCTGCACATCCAGACATGTGG 1058  
QY 981 CATGGGCTACAACTTTCAGCTTGTGGTCTGCTGGAGATCATCTTACATTTGTG 1035  
Db 1059 CATTAACAATCTGATGGCATCTCTGGCATATAGCGCGCATCTGGTGGTCTGTG 1113

RESULT 9

ABK12334 standard; cDNA; 1679 BP.

ABK12334;

05-JUN-2002 (first entry)

cDNA encoding mouse nonerythroid Rh glycoprotein Rhbg.

Mouse; nonerythroid Rh glycoprotein; Rhbg; Rhb; antibody; immunogen;

Rh type B transporter gene; polytypic transporter-type protein;

Ion transporter; chromosome 3; gene; ss.

Mus sp.

Location/Qualifiers

39..1531

/\*tag= a

/product= "Mouse Rhbg protein"

1634..1639

/\*tag= b

/standard\_name= "Polya signal"

/note= "Atypical polyadenylation site"

MO200216396-A2.

28-FEB-2002.

17-AUG-2001; 2001WO-US25881.

21-AUG-2000; 2000US-226767P.

(NYBL-) NEW YORK BLOOD CENT INC.

Huang C, Liu Z;

WPI; 2002-280905/32.

P-PSDB; AAU78092.

Homologues of mouse or human nonerythroid Rh glycoproteins, Rhbg or RhbG,

respectively, useful in the production of antibodies which are useful

for detecting Rhbg or RhbG glycoproteins in a sample -

Claim 1; Fig 1a; 59pp; English.

The present invention relates to a new protein or peptide comprising an

CC nonerythroid Rh glycoprotein homologue (Rhbg) amino acids, fully defined  
CC in the specification. The antibody of the invention is useful for  
CC detecting an Rhbg or a Rhbg glycoprotein in a sample, by contacting the  
CC sample with antibody under conditions suitable for binding, assessing the  
CC specific binding to the antibody, and thus detecting the presence of an  
CC epitope of Rhbg or RhbG in the sample. The nucleic acids of the invention  
CC are useful as probes for detecting transporter genes and particularly Rh  
CC type B transporter genes including e.g. NH<sub>4</sub><sup>4+</sup> ion transporters. The  
CC present nucleic acid sequence is that of the mouse Rhbg gene located on  
CC chromosome 3. This sequence encodes the mouse Rhbg protein of the  
CC invention. Rhbg is a polytypic transporter-type protein.

Sequence 1679 BP; 324 A; 491 C; 469 G; 395 T; 0 other;

Query Match 12.7%; Score 159.4; DB 24; Length 1679;  
Best Local Similarity 48.3%; Pred. No. 1.4e-35;  
Matches 445; Conservative 0; Mismatches 4/6; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAGATCTGACCGTATGAGCGGCGCATTTGGCTTCTTCACCTC 200  
Db 209 CTACCAAGCTTCCAGATGTCAGTCTGCTGCGCTTGGCTTGGCTTCTTCATAGT 268  
QY 201 GAGTTCCGGAGACACAGCTGAGCAGTGTGCTTCAACTCTTCATGCTGGCCTTGG 260  
Db 269 CTTCCTACAGCGGTACGCGCTTACAGCAGTGTGCTTCACTTCCCTGCGCAGCTCAC 328  
QY 261 TGTGAGTGGGCAATCCCTGTCGAGCGGCTTCTGAGCCAGTTCCTTCCGGAGAGCTGT 320  
Db 329 CTTGAGTGGGCAACTGCTGTCACAGGCTTCTTCTTCCACAGGCTGTCACATCCA 388  
QY 321 CATCACTGTTCACTATTCGCTGAGCAGCAGTATGCTTGTGCTGCTGATCTCAGT 380  
Db 389 TGTGGCGGTGAGAGATTGATCAACGCTGACTTTCGCGGAGAGCTGTCATCTCTT 448  
QY 381 GGATCTCTTCTGGGAGAGCTCACTTGGCCAGTGTGTGATGCTGCTGGTGAAGT 440  
Db 449 CGGGCTCTTCTGCGCAAGACTGGGCGCCAGCTGCTGCTAAAGGCTCACTGAGGC 508  
QY 441 GACAGCTTTAGGCAACTGAGAGATGCTATCATGATATCTTCAACACAGACTACCAT 500  
Db 509 AGTGTGTTTACGTCACAGATTTATCTACTAGTCTCTGGGGGTGAGAGATGCTGG 568  
QY 501 GAACATGATGACATCTACGTGTTCGACGCTATTTGGCTGTCTGTGCTGCTGCT 560  
Db 569 AGGTTCATGACATATTCACACTTTGGGGCTTCTGCGGCTTCTCTCAAGGCTCT 628  
QY 561 GCCAAGGCTTACCCGAGGGAACGAGAGATTAAGATACAGACGACAGATACCAATT 620  
Db 629 CTACAGATCCAGCTGAGAGAGAGCAGCATGCGCAGAGCTCTCTCAACTCTGACT 688  
QY 621 GTCTGCATGCTGGGGGCGCTCTTCTGTGATGATGTGGCCAACTTCAACTGCTGCT 680  
Db 689 CTTCGCATGATGGGACCATCTCTCTGTGGGTTTGTGGCCACTTCACTTCCGCGC 748  
QY 681 GCTGAGAGTCAATCGAAGAGAGATGCGTGTTCACACACCTATGCTGTAGAGT 740  
Db 749 GACAGCGCTGGGGATGGGAGCAGCATGGAGCGTGTCAACATCTATTCACTACCGC 808  
QY 741 CAGCGTGTGACAGCATCTCAGGTCTATCTTGGCTCAACCCCAAGGAGAGATGACAA 800  
Db 809 AAGCACCCTCAGTACTTTCCTGCTTGTGACCCCTTGTCAGTGAGTAGACGACTGACAT 868  
QY 801 GACTTATGTGCACACTGCGGTGTGGCAGAGGCGTGGCTGGTACTGCTGCTGCT 860  
Db 869 GGTCCAGCTCCAGACGACACTGTGGAGGTGTTGTGGTGGGACATCAAGTAGAT 928  
QY 861 GATCCCTTCTCCGCTTGGCATGCTGCTGCTGCTTGTGGCTGGCTGATCTCCGTCGG 920  
Db 929 GATGCTGACACCCCTTGGGGCGCTGGCAGCTGCTTCTGCTGGCTGGAGCTCTCACACT 988  
QY 921 GGGAGCCAGTACCTGCGGGGTGTGTAAACGAGTGTGGGATTTCCCAACAGCTCAT 980  
Db 989 GGGGTATTAAGTCTTTTACGCTTATCTTGAATTCAGATTTAAACGCAAGACACATGTGG 1048

QY 981 CATGGGCTACACCTTCAGCTTCGTGCTGCTTGAGAGATCATCTACATTGCTGCT 1040  
 Db 1049 TGTTCACACGTCGATCATGATGCGAGGGTCTCTGGGGCGGCATCCTCGAGTGTATGGC 1108  
 QY 1041 GGTGCTGTATACCGTCGGAGC 1061  
 Db 1109 TGCACGTGGCCACCACGAGAGC 1129  
 RESULT 10  
 AAH18728/5  
 ID AAH18728 standard; cDNA: 3706 BP.  
 AC AAH18728;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:19002.  
 XX  
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 WPI: 2001-318749/34.  
 DR  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8: SEQ ID 19002; 2537bp + CD ROM; English.  
 XX  
 XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to a  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences: AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX

SQ Sequence 3706 BP; 883 A; 1050 C; 684 G; 1089 T; 0 other;  
 Query Match 12.2%; Score 152.4; DB 22; Length 3706;  
 Best Local Similarity 99.4%; Pred. No. 2,4e-33;  
 Matches 153; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 333 CAGTATTCGGCTGGCCACCATGATGCTTTGTCGCTGCTGATCTCAGTGCATGCTGCT 392  
 Db 2906 CAGTATTCGGCTGGCCACCATGATGCTTTGTCGCTGCTGATCTCAGTGCATGCTGCT 2847  
 QY 393 GGGGAAGTCACCTTGCGCCACATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 452  
 Db 2846 GGGGAAGTCACCTTGCGCCACATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 2787  
 QY 453 CAACCTGAGATGTCATCAGTAAATATCTTCAAC 486  
 Db 2786 CAACCTGAGATGTCATCAGTAAATATCTTCAAC 2753  
 RESULT 11  
 ABR49216  
 ID ABR49216 standard; cDNA: 2098 BP.  
 XX  
 AC ABR49216;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE cDNA encoding mouse Rh type C gene (rhcg) protein.  
 XX  
 KW RHCG: mouse; non-erythroid Rh type C glycoprotein;  
 KW chromosome 7; gene; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 123..1496  
 FT FT /\*tag= a  
 FT /\*product= "rhcg protein"  
 FT polyA\_signal 1934..1939  
 FT FT /\*tag= b  
 XX  
 PN WO200220719-A2.  
 XX  
 PD 14-MAR-2002.  
 XX  
 PF 05-SEP-2001; 2001MO-US27503.  
 XX  
 PR 07-SEP-2000; 2000US-230660P.  
 XX  
 PA (NYBL-) NEW YORK BLOOD CENT INC.  
 XX  
 PI Huang C, Liu Z;  
 XX  
 DR WPI: 2002-351774/38.  
 DR P-PSDB: AAU78998.  
 XX  
 PT Nucleic acid sequences encoding novel mammalian nonerythroid Rh type C  
 PT and glycoproteins which have a characteristic twelve transmembrane  
 PT domain structure -  
 XX  
 PS Claim 3: Fig 1; 53pp; English.  
 XX  
 XX This invention relates to the nucleic acid and protein sequences of  
 CC novel human and mouse non-erythroid Rh type C glycoprotein (rhcg).  
 CC The rhcg protein and the mouse homologue (rhcg) have a characteristic  
 CC 12 transmembrane domain structure and are expressed in kidneys and  
 CC testis. The invention also comprises a method for antibody that  
 CC specifically binds an epitope of the glycoprotein and a method for  
 CC detecting the protein using this antibody. The antibodies of the  
 CC invention may be used in Western blots, enzyme linked immunosorbent  
 CC assays (ELISA) or immunohistochemical assays to identify the non-  
 CC erythroid tissues, particularly kidney and testis, that express the  
 CC rhcg or rhcg glycoproteins. The methods are used for detecting an

CC Rhog or and Rhcg glycoprotein in a sample. The present sequence  
CC represents the cDNA encoding the mouse Rhcg Rh type C glycoprotein  
CC (Rhcg) protein sequence of the invention. The gene encoding this  
CC protein is located on murine chromosome 7.

Sequence 2098 BP; 458 A; 594 C; 559 G; 487 T; 0 other;

Query Match	12.0%	Score 150.2	DB 24	Length 2098
Best Local Similarity	47.5%	Pred. No. 7.7e-33		
Matches 446	Conservative	0	Mismatches 493	Indels 0
				Gaps 0

OY	97	CACATATGACGCTCTTAAAGGATCAAAAGGGGCTCGGACATCTCTATCAAGTTGGCCAA	156
Db	258	CGCAAGAACATCTTCCAGGACGCTTGGAGACAGATTTTACTATCGCTCCGAGCTTCCAG	317
OY	157	GATCTGACCCGTGATGGCGGCATTGGGCTTGGGCTCTCCACTCCGAGTTTCCGAGACAC	216
Db	318	GATGTACACGCGCATGTCCTTGCTGGGCTTGCGGCTTCCCTCAGACCTTCTGACAGCGTAC	377
OY	217	AGCTGGAGCACTGTGGCCTTCAACCTCTTCACTGCTGGGCGCTTGGTGTGCACTGGGCAATC	276
Db	378	GGCTTCAGCGCTGAGGCTTCACCTCTCGCTGGCAGCTTTCGGCATCCACTGGGGCACTG	437
OY	277	TTGCTGGACGCGCTCTGAGCCACTTTCCTCTGGGAAGGAGGGCATCACACTGTTCAGT	336
Db	438	CTCATGCAAGGATGTTCCATTACTTTGAAAGGCCACATTTGCTTCGAGCGTGGAGAC	497
OY	337	ATTGCGGTGGCCACCATAGTGTCTTTGTGCGTGTGATCTTCAGTGAATGCTGTCTTGGAG	396
Db	498	ATCATCCAGTGACTTCTGTGTGGCATCTTCTGTGAGGCTTTCGGGGGCAAGTTCTTAGCG	557
OY	397	AAGGTCAACTTGGGGCACTTGGTGTGTATGGTGGTGGTGGAGGGTGCACAGCTTTAGGCAC	456
Db	558	AAGGTCAAGCCGATGACAGCTGCTATTATGACCTTCTTCCAAAGTACTCTTTCACAGTG	617
OY	457	CTGAGATGGTCATCAGTAATATCTTCAACACAGACTACACATGAACTGATGCATGC	516
Db	618	AATAGTTCAATCCCTCGAAGCTGATAGAGCCAAAGATGCAAGGGGGCTCTATGACCATC	677
OY	517	TACGTGTTCCGAGCTATTTTGGGCTGTCTGTGGCCTGTGGCCCTGCCAAAGCCTCTACCC	576
Db	678	CACCATTTTGGCGGCTACTTGTGGGCTCACAATGACATGACATCTCCACGAAAAAACC	737
OY	577	GAGGGAACGGAGGATAAAGATCAACAACAGCAACGATACCACATTTGTCTGCCATGCTGGCC	636
Db	738	GATCAGAGCAAGCAGAGACAGAGCTCAAGTGTACACACTCGGACCTTTTCGCCATGATTGGC	797
OY	637	GCCCTTTCTGTGATGTCTGTGGCCAAAGTTTCAACTGCTGTGCTGAGAAATGCCAATC	696
Db	798	ACCGTCTTCTGTGATATATACTGGCCCAAGTTTCAATTAGCCAGTCTCTTCCAGGGAGAT	857
OY	697	GAAAGAAAGAAATGCCGTGTTCAACACCTACTATGCTGTAGCAGTCAAGCTGATGACGCC	756
Db	858	GCCCAAGACAGGACAGCCCTCAATAACTCTCTCTGTGGAGGAGTGTCTTAACACA	917
OY	757	ATCTCAGGAGTATCCTTGGGCTCACCCCCAAGGAAGATCACAGAGCTTATGTGACAGT	816
Db	918	GTGACAGTATCAATATTGTACACAGAGGGCAAGTTGGATATGTGTGCACATCCAGAT	977
OY	817	GCGGTGTGGCAGGAGCGTGGCTGTGGGTAACCTGATGTCACTGATCACTGATCCCTTCCCGTGG	876
Db	978	GCCACGCTTGCAGGTGGGGTGGTGTGGGCAACAGCTGGGAGATGATGTCACACCTTAC	1037
OY	877	CTTGTCCATGGTGTGGGCTGTGTGGCTGTGGCTGATATCTCCGTGGGGGAGCCAAATGACTG	936
Db	1038	GGCGCTCTCATCGTGGGGTCTTCTGTGGGCAATTTTCTCCACCCAGGATTTTGCATACCTA	1097
OY	937	CCGGGGTGTGTACCAAGATGCTGGGGATTTCCCAAGACTCCATCATGGGCTACAACTTC	996
Db	1098	ACGCGATTTCCGTGAAGTCCCGGCTTCGCATCCAGGACACATGTGGCATTCACAACCTGCAC	1157
OY	997	AGCTTGTGGGCTGTGGAGAGATCATCTACATGTGTG	1035

Db 1158 GGCATTCTGGCATCATAGCGGCATTGTGGTCTGTG 1196

RESULT 12  
ABA42886/c  
ID ABA42886 standard; DNA; 486 BP.

AC	ABA42886;
XX	
DT	01-FEB-2002 (first entry)

DE Human breast cell single exon nucleic acid probe #1581.

KW Human; microarray; single exon probe; gene expression; breast; disease; cancer; ss.

OS Homo sapiens.

PN WO200157271-A2

PD 09-AUG-2001

PF 30-JAN-2001; 2001WO-US006662.  
YY

PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456

PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.

PR	21-SEP-2000; 2000US-0234067.
PR	27-SEP-2000; 2000US-0236359.

XX  
XX

XX  
XX

XX

XX

useful for measuring gene expression

XX  
XX  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840

xx  
Ebo

CC nucleic acid probes for measuring gene expression in a sample derived from human breast and BT 474 cells. The method involves contacting

the probes with a collection of detectably labelled nucleic acids derived from mRNA of human breast, and then measuring the

CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a single exon nucleic acid probe of the invention  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

SQ. Sequence 486 BP; 137 A; 112 C; 112 G; 125 T; 0 other;

Query Match	11.1%	Score 139.8	DB 22	Length 486
Best Local Similarity	97.5%	Pred No. 3.7e-30		
Matches 147	Conservative 0	Mismatches 12	Indels 0	Gaps 0

Dy 477 TATCTTCACACGAGACTACCACATGTAACTGACATCTCAGTGTTCCGAGCGTATT 536  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 458 TTATTTGCAGACGACTACCAACATGAACCTGAGGCACTTCACTGTTGCCAGCGCTATT 399

QY 537 TGGGCTGTCTGTGCTGCTGCTGCCAAGCCTTACCCGAGGGAAGGAGATTAAGA 596  
|||||  
Db 398 TGGGCTGTGACTGTGCTGCTGCTGCCAAGCCTTACCCGAGGGAAGGAGATTAAGA 339  
597 TCAGACAGCAAGCATACCCAGTTGTCTGCCATGCTGGG 635  
Db 338 TCAGACAGCAAGCATACCCAGTTGTCTGCCATGCTGGG 300

## RESULT 13

ABAB3307/c  
ID ABA53307 standard; DNA; 486 BP.

AC ABA53307;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #1612.

DE Human: foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PT WPI; 2001-483447/52.

PS Claim 1; SEQ ID NO 1612; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 486 BP; 137 A; 112 C; 112 G; 125 T; 0 other;

Query Match 11.1%; Score 139.8; DB 22; Length 486;  
Best Local Similarity 92.5%; Pred. No. 3.7e-30;  
Matches 147; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 477 TATCTCAACAGCACTACACATGAACATGATGACACATCTAGTGTGCGACCTATT 536  
|||||  
Db 458 TTATTATGACAGCACTACACATGAACCTGAGGCACTTCTACTGCTTGGACCTATT 399

QY 537 TGGGCTGTCTGTGCTGCTGCCAAGCCTTACCCGAGGGAAGGAGATTAAGA 596  
|||||  
Db 398 TGGGCTGTGACTGTGCTGCTGCCAAGCCTTACCCGAGGGAAGGAGATTAAGA 339

QY 597 TCAGACAGCAAGCATACCCAGTTGTCTGCCATGCTGGG 635

Db 338 TCAGACAGCAAGCATACCCAGTTGTCTGCCATGCTGGG 300  
|||||

## RESULT 14

ABAB3081/c  
ID ABA23081 standard; DNA; 486 BP.

AC ABA23081;

DT 23-JAN-2002 (first entry)

DE Probe #1547 for gene expression analysis in human heart cell sample.

DE Human: gene expression; heart; microarray; vascular system; probe;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease; ss.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PT WPI; 2001-48899/53.

PS Claim 1; SEQ ID NO 1547; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart. The

CC present sequence is one such probe. The probes may be used for

CC predicting, measuring and displaying gene expression in samples derived

CC from the human heart via microarrays. By measuring gene expression, the

CC probes are useful for predicting, diagnosing, grading, staging,

CC monitoring and prognosing diseases of the human heart and vascular system

CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and

CC congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pcl\_sequences.

Query Match 11.1%; Score 139.8; DB 22; Length 486;  
Best Local Similarity 92.5%; Pred. No. 3.7e-30;  
Matches 147; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 477 TATCTCAACAGCACTACACATGAACATGATGACACATCTAGTGTGCGACCTATT 536  
|||||  
Db 458 TTATTATGACAGCACTACACATGAACCTGAGGCACTTCTACTGCTTGGACCTATT 399

QY 537 TGGGCTGTCTGTGCTGCTGCCAAGCCTTACCCGAGGGAAGGAGATTAAGA 596  
|||||  
Db 398 TGGGCTGTGACTGTGCTGCTGCCAAGCCTTACCCGAGGGAAGGAGATTAAGA 339

QY 597 TCAGACAGCAAGCATACCCAGTTGTCTGCCATGCTGGG 635



THIS PAGE BLANK (USPTO)





Db 91 GAAGCAGCTCTCATCTCTCTCTATTTTAACTACCATGACGCTTCTTAGAGAT 150  
QY 121 CAAAGGGGCTGCTGGCATCTATCAAGTTGGCCAAAGATCTGACCGTATGGGCGCAT 180  
Db 151 CAAAGGGGCTGCTGGCATCTATCAAGTTGGCCAAAGATCTGACCGTATGGGCGCAT 210  
QY 181 GCGTTGGGCTCTCTACCTGAGTTTCCGAGACACAGCTGGAGCACTGTGGCTTCAAC 240  
Db 211 GCGTTGGGCTCTCTACCTGAGTTTCCGAGACACAGCTGGAGCACTGTGGCTTCAAC 270  
QY 241 CTCTCATGCTGGCGCTTGGTGTGACAGTGGCAATCTGCTGGACGCTTCTGAGCAG 300  
Db 271 CTCTCATGCTGGCGCTTGGTGTGACAGTGGCAATCTGCTGGACGCTTCTGAGCAG 330  
QY 301 TTCCCTTCTGGAGAGTGTATCACACTGTTCAATTCGGCTGGCCACCATGAGTGT 360  
Db 331 TTCCCTTCTGGAGAGTGTATCACACTGTTCAATTCGGCTGGCCACCATGAGTGT 390  
QY 361 TTGTCGGTGTATCTCAGTGTGATGCTGTTGGGGAAGTCAACTTGGCGCAGTTGGTG 420  
Db 391 ATGTCGGTGTATCTCAGTGTGATGCTGTTGGGGAAGTCAACTTGGCGCAGTTGGTG 450  
QY 421 GTATGAGTGTGGAGTGTGACAGCTTTAGGCACTGAGGATGGTCACTAGTAATATC 480  
Db 451 GTATGAGTGTGGAGTGTGACAGCTTTAGGCACTGAGGATGGTCACTAGTAATATC 510  
QY 481 TTCAACACAGACTACACATGAACTGATGATGACATCTACGTTTTCGAGCCTATTTGGG 540  
Db 511 TTCAACACAGACTACACATGAACTGATGATGACATCTACGTTTTCGAGCCTATTTGGG 570  
QY 541 CTGTCGTGTGGCTGTGGTGTGCTGCAAGCCCTTACCCGAGGAACGAGATTAAGATCAG 600  
Db 571 CTGTCGTGTGGCTGTGGTGTGCTGCAAGCCCTTACCCGAGGAACGAGATTAAGATCAG 630  
QY 601 ACAGCAGACTACCCAGTTTGTCTGCAATGCTGGGCGCTTCTTGTGATGATGCTG 660  
Db 631 AGAGCAAGATACCCAGTTTGTCTGCAATGCTGGGCGCTTCTTGTGATGATGCTG 690  
QY 661 CCAAGTTTCACTGCTGTCTGTGAGAAGTCCAAATGGAAGAAAGATGCGGTTCAC 720  
Db 691 CCAAGTTTCACTGCTGTCTGTGAGAAGTCCAAATGGAAGAAAGATGCGGTTCAC 750  
QY 721 ACCTACTATGCTTACAGTCAAGCTGTGAGAGCCATCTCAGGGTCTCTTGGCTCAC 780  
Db 751 ACCTACTATGCTTACAGTCAAGCTGTGAGAGCCATCTCAGGGTCTCTTGGCTCAC 810  
QY 781 CCCCAGGGAATACAGCAAGACTTATGTCACAGTGGCGGTGTGGCAGAGGCGTGGCT 840  
Db 811 CCCCAGGGAATACAGCAAGACTTATGTCACAGTGGCGGTGTGGCAGAGGCGTGGCT 870  
QY 841 GTGGGTACCTGTGTACCTGATCCCTTCCGTGGCTTGGCATGTTGCTGGGTCTTGTG 900  
Db 871 GTGGGTACCTGTGTACCTGATCCCTTCCGTGGCTTGGCATGTTGCTGGGTCTTGTG 930  
QY 901 GCTGGGCTGATCTCCGTGGGGAAGCAAGTACCTGCGGGGTGTGTAAACGAGTGTG 960  
Db 931 GCTGGGCTGATCTCCGTGGGGAAGCAAGTACCTGCGGGGTGTGTAAACGAGTGTG 990  
QY 961 GGGATTCCCAAGCTACATGAGGCTCAACTTACGTTGCTGGGTCTGTGGAGAG 1020  
Db 991 GGGATTCCCAAGCTACATGAGGCTCAACTTACGTTGCTGGGTCTGTGGAGAG 1050  
QY 1021 ATCATCTACATTTGCTGTGTGTGTGATACCGTGGAGCCGCAATGGCATGATGGC 1080  
Db 1051 ATCATCTACATTTGCTGTGTGTGTGATACCGTGGAGCCGCAATGGCATGATGGC 1110  
QY 1081 TTCCAGGTCTCTCAGCATTTGGGGAATCAGCTTGGCAATCTGATAGCTCAGCTCT 1140  
Db 1111 TTCCAGGTCTCTCAGCATTTGGGGAATCAGCTTGGCAATCTGATAGCTCAGCTCT 1170  
QY 1141 GGTCTCTGACAGCTTTGGTCCCTAAATCTTAAATATGGAAGACACCTCATGGGCTTAA 1200  
Db 1171 GGTCTCTGACAGCTTTGGTCCCTAAATCTTAAATATGGAAGACACCTCATGGGCTTAA 1230

QY 1201 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1254  
Db 1231 TATTTGATGACCAAGTTTCTGGAAGTTTCTCATTTGGCTGTGGATTTTAA 1284  
RESULT 2  
US-08-553-888A-2  
Sequence 2, Application US/08553888A  
Patent No. 5723293  
GENERAL INFORMATION:  
APPLICANT: Huang  
TITLE OF INVENTION: DIAGNOSTIC METHOD AND KIT FOR  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann & Baron, LLP  
STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: USA  
ZIP: 11753  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/553, 888A  
FILING DATE: 11/06/95  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: O'Dea, Sean W.  
REGISTRATION NUMBER: 37690  
REFERENCE/DOCKET NUMBER: 454-5  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
INFORMATION FOR SEO ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1466 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-553-888A-2  
Query Match 94.3%; Score 1182; DB 1; Length 1466;  
Best Local Similarity 96.4%; Pred. No. 0;  
Matches 1209; Conservative 0; Mismatches 45; Indels 0; Gaps 0;  
QY 1 ATGAGCTTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTCTGGGCGCTTAACATG 60  
Db 45 ATGAGCTTAAGTACCCGCGGTCTGTCCGCGCTGCTGCCCTCTGGGCGCTTAACATG 104  
QY 61 GAAGCAGCTCTATTTCTCTCTTCTATTTTAACTACCATGACGCTTCTTAGAGAT 120  
Db 105 GAAGCAGCTCTATTTCTCTCTTCTATTTTAACTACCATGACGCTTCTTAGAGAT 164  
QY 121 CAAAGGGGCTGCTGGCATCTATCAAGTTGGCCAAAGATCTGACCGTATGGGCGCAT 180  
Db 165 CAAAGGGGCTGCTGGCATCTATCAAGTTGGCCAAAGATCTGACCGTATGGGCGCAT 224  
QY 181 GCGTTGGGCTCTCTACCTGAGTTTCCGAGACACAGCTGGAGCACTGTGGCTTCAAC 240  
Db 225 GCGTTGGGCTCTCTACCTGAGTTTCCGAGACACAGCTGGAGCACTGTGGCTTCAAC 284  
QY 241 CTCTTATGCTGGCGCTTGGTGTGACAGTGGCAATCTGCTGGACGCTTCTGAGCAG 300  
Db 285 CTCTTATGCTGGCGCTTGGTGTGACAGTGGCAATCTGCTGGACGCTTCTGAGCAG 344  
QY 301 TTCCCTTGGGAAGTGTGATCACAGTGTTCAGTATTCGGCTGGCCACCATGAGTGTCT 360  
Db 345 TTCCCTTGGGAAGTGTGATCACAGTGTTCAGTATTCGGCTGGCCACCATGAGTGTCT 404

QY 361 TTGTGGTGTGATCTCAGTGTGATCTGTCTTGGGGAAGTCAACTTGGCGCAGTGGTG 420  
 Db 405 ATGTGCGGTGATCTCAGCGGGTGTCTCTTGGGGAAGTCAACTTGGCGCAGTGGTG 464  
 QY 421 GTGATGCTGTGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGTGATCATGATATATC 480  
 Db 465 GTGATGCTGTGTGGAGGTGACAGCTTTAGGCAACCTGAGGATGTGATCATGATATATC 524  
 QY 481 TTCAACACAGCTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
 Db 525 TTCAACACAGCTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 584  
 QY 541 CTGTGTGCTG 600  
 Db 585 CTGTGTGCTG 644  
 QY 601 ACAGCAAGATATCCAGTG 660  
 Db 645 AGACCAAGATATCCAGTG 704  
 QY 661 CCAGTTTCACTG 720  
 Db 705 CCAAGTGTCACTG 764  
 QY 721 ACCTACTATG 780  
 Db 765 ACCTACTATG 824  
 QY 781 CCCCAGAGGAGATCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
 Db 825 CCCCAGAGGAGATCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 884  
 QY 841 GTGGGTACTG 900  
 Db 885 GTGGGTACTG 944  
 QY 901 GCTGGGTACTG 960  
 Db 945 GCTGGGTACTG 1004  
 QY 961 GGGATTTCCCAACAGCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 Db 1005 GGGATTTCCCAACAGCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1064  
 QY 1021 ATCACTTCACTG 1080  
 Db 1065 ATCACTTCACTG 1124  
 QY 1081 TTCCAGTG 1140  
 Db 1125 TTCCAGTG 1184  
 QY 1141 GGTCTCTG 1200  
 Db 1185 GGTCTCTG 1244  
 QY 1201 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTGATTTGGCGTGTGTGTGTAA 1254  
 Db 1245 TATTTTGTATGACCAAGTTTCTGGAAGTTTCTGATTTGGCGTGTGTGTGTAA 1298

## RESULT 3

US-09-461-325-75  
 ; Sequence 75, Application US/09461325A  
 ; Patent No. 6475753  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et al.  
 ; TITLE OF INVENTION: 94 Human Secreted Proteins  
 ; FILE REFERENCE: P2029pl  
 ; CURRENT APPLICATION NUMBER: US/09/461, 325A  
 ; CURRENT FILING DATE: 1999-12-14  
 ; EARLIER APPLICATION NUMBER: PCT/US99/13418

; EARLIER FILING DATE: 1999-06-15  
 ; EARLIER APPLICATION NUMBER: 60/089, 507  
 ; EARLIER FILING DATE: 1998-06-16  
 ; EARLIER APPLICATION NUMBER: 60/089, 508  
 ; EARLIER FILING DATE: 1998-06-16  
 ; EARLIER APPLICATION NUMBER: 60/089, 509  
 ; EARLIER FILING DATE: 1998-06-16  
 ; EARLIER APPLICATION NUMBER: 60/089, 510  
 ; EARLIER FILING DATE: 1998-06-16  
 ; EARLIER APPLICATION NUMBER: 60/090, 112  
 ; EARLIER FILING DATE: 1998-06-22  
 ; EARLIER APPLICATION NUMBER: 60/090, 113  
 ; EARLIER FILING DATE: 1998-06-22  
 ; NUMBER OF SEQ ID NOS: 532  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 75  
 ; LENGTH: 1650  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-461-325-75

Query Match 8.7%; Score 109.6; DB 4; Length 1650;  
 Best Local Similarity 48.7%; Pred. No. 4,5e-23;  
 Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1;

QY 364 TCGGTGTGATCTCAGTGTGATCTGTCTTGGGGAAGTCAACTTGGCGCAGTGGTG 423  
 Db 33 TCTGTCTCCTG 92  
 QY 424 ATG 483  
 Db 93 ATGACTTCTTCTTCCAGTG 152  
 QY 484 AACACAGCTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 543  
 Db 153 AAGGTGAAGGATGACGAGAGGCTTCCATGACATCCACATTTGGGCGCTACTTGGGCTC 212  
 QY 544 TCTG 603  
 Db 213 ACAGTG 272  
 QY 604 GCACGATATCCAGTTTG 663  
 Db 273 GTGTACAGAGTG 332  
 QY 664 AGTTCAACTG 723  
 Db 333 AGCTTCAACTG 392  
 QY 724 TACTATG 783  
 Db 393 TACTATG 451  
 QY 784 CAAGGAGATGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 843  
 Db 452 AAGGCAAGCTG 511  
 QY 844 GGTACTG 903  
 Db 512 GGTACTG 571  
 QY 904 GGGGTG 963  
 Db 572 GGCATATCTTCCAGCTG 631  
 QY 964 ATTCCCAACAGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1023  
 Db 632 ATCCAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 691  
 QY 1024 ATCTACATTTGTG 1035  
 Db 692 GTGGGTGTGTGTG 703



QY	950	ACCGAGTGTGGGGATTCCCCACAGCTCATAGGGCTACAACTTACGTTGCTGGGTC	1009
Db	1677	GCAGCAACCCGCCCCAAAGCGTCGCCCGCGGGGCCGCTGGCAGCGGGTCCGCT	1736
QY	1010	TGCTTGAGAGATCATCTACATGTGCTGCTGCTTGGATACCGTGGAGACGGCAATG	1059
Db	1737	CGCTGCCCGGTTTCATCTGCTGTTGATGCTGGGGACAGACTCGGTCGCCACTACT	1796
QY	1070	GCATGATTGGGCTCCAGGTCCTCCTCAGC	1098
Db	1797	ACATGAAGGATCTGCTCCCTACCGAGGC	1825

RESULT 7  
 US-09-252-991A-8351/C  
 ; Sequence 8351, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERGININOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

? FILE REFERENCE: 107196.136
? CURRENT APPLICATION NUMBER: US/09/252,991A
? CURRENT FILING DATE: 1999-02-18
? PRIOR APPLICATION NUMBER: US 60/074,788
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/094,190
? PRIOR FILING DATE: 1998-07-27
? NUMBER OF SEQ ID NOS: 33142
? SEQ ID NO 8351
? LENGTH: 1422
? TYPE: DNA
? ORGANISM: Pseudomonas aeruginosa
? US-09-252-991A-8351

```

[illegible]

```

RESULT 8
US-09-252-991A-8488
: Sequence 8488, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252.991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142

```

```

; SEQ ID NO 8488
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8488

```

Query Match	3.2%	Score	40.6	DB	4	Length	1503
Best Local Similarly	51.4%	Pred. No.	0.039				
Matches	94	Conservative	0	Mismatches	89	Indels	0
						Gaps	0

[illegible]

RESULT 9  
US-09-252-991A-8423  
; Sequence 8423, Application US/09252991A  
; Patent No. 6551795

```

1  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
2  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
3  FILE REFERENCE:  107196.136
4  CURRENT APPLICATION NUMBER:  US/09/252,991A
5  CURRENT FILING DATE:  1999-02-18
6  PRIOR APPLICATION NUMBER:  US 60/074,788
7  PRIOR FILING DATE:  1998-02-18
8  PRIOR APPLICATION NUMBER:  US 60/094,190
9  PRIOR FILING DATE:  1998-07-27
10 NUMBER OF SEQ ID NOS:  33142
11 SEQ ID NO 8423
12
13 LENGTH: 1761
14
15 TYPE: DNA
16
17 ORGANISM:  Pseudomonas aeruginosa
18
19 US-09-252-991A-8423

```

Query Match	3.2%	Score 40.6:	DB 4:	Length 1761:
Best Local Similarity	51.4%:	Pred. No. 0.042:	Indels 0:	Gaps 0:
Matches	94:	Conservative	0:	Mismatches 89:
OY	343	CTGGCCACCAAGAGAGTCTTTGTCGGTGTGATCTCAGTGAATGCTGTCTTGGGGAAGGTC	402	
Db	712	CTGGCGAATGATCAACGCCACTGTGGCGAAGATCTACGTGAACCGAATTTGTCCCCAAGGTC	771	
OY	403	AACCTTGGCGCAGTTTGCTGTATGCTGCTGGTGGAGGTGAACAGCTTCTTAGGCAACTGTGAGG	462	
Db	772	GCCCCGTGGATCTGGGTGGTGTATGCTTGGCTGATCATGACCCCTCATCAACCTCAAGAGC	831	
OY	463	ATGGCATCAAGTAATATCTTCAACACACATACACATCAAGACATGATGCACATCTAAGCTG	522	
Db	832	GTGAACCTCGGTGGCCAAATTTCAACACCGCTCTTCTGCTGCTGACCTGGCGATCATCGTG	891	
OY	523	TTC	525	
Db	892	GTC	894	

RESULT 10  
US-09-252-991A-6902/c  
; Sequence 6902, Application US/09252991A

Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074.788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094.190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 6902  
LENGTH: 801  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-6902

Query Match 3.2%; Score 40; DB 4; Length 801;  
Best Local Similarity 47.9%; Pred. No. 0.04;  
Matches 115; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 236 TCACCTCTTCATGCTGGCCCTTGCTGTGACGTGGCAATCCTGCTGGAGGCTTCTGA 295  
DB 742 TCCTGCTCTGGCCCTGCACTTCTGCTGCGCGGATCAAGAGCGCGCTTCATCA 683  
QY 296 GCCAGTTCCTTGGGAGGTGTCATCAGTTCAGTATTCGGCTGGCCACCATGA 355  
DB 682 ACACCGTACACCGGTGGCCAGGTGTCGCGCTGTCCTGTCATCTGCTGT 623  
QY 356 GTGCTTGTGCTGCTGATCTCAGTGAGTCTCTTGGGGAAGTCACTTGGCGCAGT 415  
DB 622 TCCCTTCAAGCTGACATCTTCACCGCCGACATCTGGGCGCAGACACCCGACCTGG 563  
QY 416 TGGTGGTATGCTGTGGAGGTGACACCTTTAGGCAACCTGAGATGCTATCAGTA 475  
DB 562 GCACCGTATGAACCAAGTGGCCGCAACATGCTGTGTCACCGTGGTTCATCGGCA 503

RESULT 11  
US-09-252-991A-6871/C  
Sequence 6871, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074.788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094.190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 6871  
LENGTH: 1410  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-6871

Query Match 3.2%; Score 40; DB 4; Length 1410;  
Best Local Similarity 47.9%; Pred. No. 0.056;  
Matches 115; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 236 TCACCTCTTCATGCTGGCCCTTGCTGTGACGTGGCAATCCTGCTGGAGGCTTCTGA 295  
DB 1046 TCCTGCTCTGGCCCTGCACTTCTGCTGCGCGGATCAAGAGCGCGCTTCATCA 987  
QY 296 GCCAGTTCCTTGGGAGGTGTCATCAGTTCAGTATTCGGCTGGCCACCATGA 355  
DB 986 ACACCGTACACCGGTGGCCAGGTGTCGCGCTGTCCTGTCATCTGATCTGCTGT 927

QY 356 GTGCTTGTGCTGCTGATCTCAGTGAGTCTCTTGGGGAAGTCACTTGGCGCAGT 415  
DB 926 TCCCTTCAAGCTGACATCTTCACCGCCGACATCTGGGCGCAGACCAACCCGACCTGG 867  
QY 416 TGGTGGTATGCTGTGGAGGTGACACCTTTAGGCAACCTGAGATGCTATCAGTA 475  
DB 866 GCACCGTATGAACCAAGTGGCCGCAACATGATGCTGTGTCACCGTGGTTCATCGGCA 807

RESULT 12  
US-09-252-991A-6850  
Sequence 6850, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074.788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094.190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 6850  
LENGTH: 1482  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-6850

Query Match 3.2%; Score 40; DB 4; Length 1482;  
Best Local Similarity 47.9%; Pred. No. 0.058;  
Matches 115; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 236 TCACCTCTTCATGCTGGCCCTTGCTGTGACGTGGCAATCCTGCTGGAGGCTTCTGA 295  
DB 417 TCCTGCTCTGGCCCTGCACTTCTGCTGCGCGGATCAAGAGCGCGCTTCATCA 476  
QY 296 GCCAGTTCCTTGGGAGGTGTCATCAGTTCAGTATTCGGCTGGCCACCATGA 355  
DB 477 ACACCGTACACCGGTGGCCAGGTGTCGCGCTGTCCTGTCATCTGATCTGCTGT 536  
QY 356 GTGCTTGTGCTGCTGATCTCAGTGAGTCTCTTGGGGAAGTCACTTGGCGCAGT 415  
DB 537 TCCCTTCAAGCTGACATCTTCACCGCCGACATCTGGGCGCAGACCAACCCGACCTGG 596  
QY 416 TGGTGGTATGCTGTGGAGGTGACACCTTTAGGCAACCTGAGATGCTATCAGTA 475  
DB 597 GCACCGTATGAACCAAGTGGCCGCAACATGATGCTGTGTCACCGTGGTTCATCGGCA 566

RESULT 13  
US-09-252-991A-6757  
Sequence 6757, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074.788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094.190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 6757  
LENGTH: 1500  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa







QY	706	AATGCCGCTGCACACCTACTATCTGTAGCAGTACACGGTGGTACAGCCATCTCAGGG	765
Db	56	ATTCGCCGTGTTCAACACCTACTATCTGTAGCAGTACACGGTGGTACAGCCATCTCAGGG	115
QY	766	TGATCTTGGCTACACCCCAAGGGAAGATACACAAGATCTTATGTGGACATAGCGCGTATG	825
Db	116	TGATCTTGGCTACACCCCAAGGGAAGATACACAAGATCTTATGTGGACATAGCGCGTATG	175
QY	826	GCAGAGAGCGTGGCTGTGGGTACCTGTGTCACTGTATCCCTTCTCGTGGCTTGGCATG	885

Db 176 GCAGGAGGCGTGGCTGTGGGTACCTGCTACCTATCCCTTCCGTTGGCTTGCATG 235  
QY 886 GTGCTGGGTCTGTGGCTGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCGGGGTGT 945  
Db 236 GTGCTGGGTCTGTGGCTGGCTGATCTCCGTCGGGGAGCCAAAGTACCTGCGGGGTGT 295  
QY 946 TGTAAACGAGTGTGGGGATTCCCAACAGCTCCATATGAGGTACACTTTCAGCTTGTG 1005  
Db 296 TGTAAACGAGTGTGGGGATTCCCAACAGCTCCATATGAGGTACACTTTCAGCTTGTG 355  
QY 1006 GGTCTCTGGAGAGATCATATGATGTGCTGGTGTGATACCGTCGGAGCGGGC 1065  
Db 356 GGTCTCTGGAGAGATCATATGATGTGCTGGTGTGATACCGTCGGAGCGGGC 415  
QY 1066 AATGGCATG 1074  
Db 416 AATGGCATG 424

RESULT 2  
US-09-949-145-5  
; Sequence 5, Application US/0949145  
; Patent No. US20020055622A1  
; GENERAL INFORMATION:  
; APPLICANT: New York Blood Center  
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and Glyco  
; FILE REFERENCE: Docket 454-31  
; CURRENT APPLICATION NUMBER: US/09/949,145  
; CURRENT FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: US 60/230660  
; PRIOR FILING DATE: 2000-09-07  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1440  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-949-145-5

Query Match 13.2%; Score 165.4; DB 9; Length 1440;  
Best Local Similarity 49.1%; Pred. No. 2,4e-42;  
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAAAGATCTGACGCTGATGGCGGCCATTTGGCTTGGCTTCTACCTC 200  
Db 177 CTACCCAAAGCTTCCAGGAGCTGACGCTGATGGCTTCTGCGGGCTTCGCTCTCATGAC 236  
QY 201 GAGTTTCGGAGACAGCTGAGAGTGTGGCTTCAACCTTCATCATCTGCGCTTGG 260  
Db 237 TTTCTGACGCGCTACGCGCTTCAGCGCGCTGGGCTTCACTTCTGTTGGACGCTTCGG 296  
QY 261 TGTGAGTGGGCAATCTGCTGAGGAGCTTCTGAGCCAGTTCCTTGGGAAGGTGT 320  
Db 297 CATCCAGTGGGGGCTGCTATGCAAGGCGTGTTCACACTTCTTACAAGACCGCTACATCT 356  
QY 321 CATCAACTGTTCACTATTTGCGCTGGCCACCATGAGTGTCTTGTGGTGTGATCTCACT 380  
Db 357 CTGGGGCGTGGAAACCTCATCAACGCTGACTTTCGCGGCTGTGTGTGGTGGCCCTT 416  
QY 381 GGATGCTGTCTGGGGAAGGTCACATGTGGCGAGTTGTGGTGTGATGTGCTGTGAGGT 440  
Db 417 TGGGGCAGTTCTGGGTAAAGTACGCCCACTTACGCTGCATCATGACTTCTTCCAACT 476  
QY 441 GACAGCTTTAGGCAACTGAGAGTGTATCATAGTAATATCTTCAACACAGACTACACAT 500  
Db 477 GACCCCTTTCGCTGTGAATGATGATTCATTCCTTAAACGCTTAAGGTAAAGATGACAG 536  
QY 501 GAACATGATGACATCTACGTTTGGCAAGCTTATTTGGGCTGTGTGGCTGTGCTCT 560  
Db 537 AGGCTTCATGACATCCACATTTGGGCTTACTTTGGGCTCACAGTGAACCCGGATCTT 596  
QY 561 GCCAAGCCTCTAACCCGAGGAGGATGAATGAATGAGACAGACATGCCAGTTT 620

Db 597 CTACCGACGCACTAGACGACGAGAGACGAATTTGTGTACCACTGCGACCT 656  
QY 621 GTCTGCCATCTGGGGCGCCCTTCTTGTGATGTTCTGGCAAGTTTCAACTGTCTCT 680  
Db 657 CTCTGCCATCTGGGACCCCTTCTCTGTGATGTTACTTGGCCAGCTTCACTGACCCAT 716  
QY 681 GCTGAGAGTCCCAATCGAAGAGATGCCGTGTTCACACACTTACCTATGTGTAGCACT 740  
Db 717 ATCTTACCATGGGGAGACGACACAGACCCGCCATCAACACTACTGCTCTCTGGACAC 776  
QY 741 CAGCGGTGTGACAGCATCTGACAGGTATCTTGGCTCACCCCAAGGAAATACAGCA 800  
Db 777 GTGCGCTTACCTCGGTGTGCAATATCAGTCCCTTGCACAAAGAGGCAACTGTGACAT 836  
QY 801 GACTTATGTACAGAGTGGGTGTGGCAGAGCGGTGGCTGTGATCTGTCTACCT 860  
Db 837 GGTGCACTACCAAGATCCACAGCTGCAAGAGGGGTGGCGGTGTGACGCTGTGAGAT 896  
QY 861 GATCCCTTTCGCTGGCTTGCATGTGTGTGGTCTTGTGTGGCTGTGATCTGTGCG 920  
Db 897 GATGCTCATGCTTACGAGTGGCCCTCATATGAGCTTGTGCTGCGCATCATCTCCACCT 956  
QY 921 GGGAGCCAGTACCTGCCGGGTGTGTATACGAGTGTGGGATTTCCCAAGCTCCAT 980  
Db 957 GGGTTTGTATACCTGACCCCAATCTGTGAGTCCCGGTGACATTCACAGACATGTGG 1016  
QY 981 CATGGGCTACACTTACGCTTGTGCTGTGGTGTGAGAGATCATCTACATGTG 1035  
Db 1017 CATTAACATCTGCAATGCGATTCCTGTGCAATCATATGAGGCGCATGTGGTGTGTG 1071

RESULT 3  
US-09-949-145-1  
; Sequence 1, Application US/0949145  
; Patent No. US20020055622A1  
; GENERAL INFORMATION:  
; APPLICANT: New York Blood Center  
; TITLE OF INVENTION: Mammalian No. US20020055622A1-erythroid Rh Type C Genes and G1  
; FILE REFERENCE: Docket 454-31  
; CURRENT APPLICATION NUMBER: US/09/949,145  
; CURRENT FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: US 60/230660  
; PRIOR FILING DATE: 2000-09-07  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1952  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION: AF193809  
; DATABASE ACCESSION NUMBER: AF193809  
; DATABASE ENTRY DATE: 1999-12-22  
; RELEVANT RESIDUES: (1)..(1952)  
US-09-949-145-1

Query Match 13.2%; Score 165.4; DB 9; Length 1952;  
Best Local Similarity 49.1%; Pred. No. 2,9e-42;  
Matches 439; Conservative 0; Mismatches 456; Indels 0; Gaps 0;

QY 141 CTATCAAGTTGGCCAAAGATCTGACGCTGATGGCGGCCATTTGGCTTGGCTTCTACCTC 200  
Db 201 CTACCCAAAGCTTCCAGGAGCTGACGCTGATGTCTTCTGGGCTTCCGCTTCTCATAGC 260  
QY 201 GAGTTTCGGAGACACAGCTGAGAGTGTGGCTTCAACCTTTCATGCTGGCGCTTGG 260  
Db 261 TTTCTGACGCGCTACGCGCTTCAGCGCGTGGGCTTCAACTTCTGTTGGACGCTTGG 320  
QY 261 TGTGAGTGGGCAATCTGCTGAGAGCGCTTCTGTAGCCAGTTCCTTGTGGGAAGGTGT 320  
Db 321 CATCCAGTGGGGGCTGCTATGACAGGCTGTGTTCACTTCTTACAAGACCCCTACATCT 380  
QY 321 CATCAACTGTTCACTATTTGCGCTGGCCACCATGAGTGTCTTGTGGTGTGATCTCACT 380



```

; PRIOR APPLICATION NUMBER: US 60/230660
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2097
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATON INFORMATION:
; DATABASE ACCESSION NUMBER: AF193810
; DATABASE ENTRY DATE: 1999-12-22
; RELEVANT RESIDUES: (1)..(2097)
US-09-949-145-2

```

```

Query Match      12.0%; Score 150.2; DB 9; Length 2097;
Best Local Similarity 47.5%; Pred. No. 2.4e-37;
Matches 446; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

QY 97 CACTATGACGCTTCCCTTAAGAGATCAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAA 156
DB 258 CCGAAGAACATCTCCAGCGACGTTGAAAGAGATTACTATCGCTACCCGAGCTTCCAG 317
QY 157 GATCTGACCGCTGATGGCGGCATGTGGCTTCCGACCTCGAGTTTCCGGAGACAC 216
DB 318 GATGTAACAGCCATGCTTCTGCGCTTCCGCTTCTCATGACCTTCTCGAGCGCTAC 377
QY 217 AGCTGAGAGAGTGTGGCTTCAACCTCTTCATGTGGCGCTTGTGTGCGAGTGGGCAATC 276
DB 378 GGCTTCAGCGGTGTAGGCTTCAACTCCGTGTGGAGCTTTGGGATCCAGAGGGGACACG 437
QY 277 CTGCTGAGAGGCTTCTGAGCCAGTTCCCTTGTGGAGAGTGTGATCATCACTGTTCACT 336
DB 438 CTCATGACAGGAGATGTTCCATTACTTTGAAGAAGGCCAATGTCTCTGAGCGTCGAGAAC 497
QY 337 ATTGGCTGGCCACCATGATGCTTGTGGTGTGATCTCATGTGATGCTGTGCTTGGGG 396
DB 498 ATCATTCAGAGTACTTCTGTGTGGCATTTCTCTGTGTGGCTTCCGGGAGTTTATAGGC 557
QY 397 AAGGTCAACTTGGCGCAGTTGGTGTGATGTGTGGTGGAGAGTGAACACTTATAGCAAC 456
DB 558 AAGGTGAGCCGATGACGCTCATATTAAGCTTCTTCAAGTACCTCTTTCACAGTGC 617
QY 457 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 516
DB 618 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 677
QY 517 TACGTTTGGCAGCAGCTATTTGGGCTGTGTGGCTGTGGCTGTGGCAAGCCTTACCC 576
DB 678 CACACATTTGGGCGCTACTTTGGGCTCACAGTACCTGATCTCTACCGAAAAACCTG 737
QY 577 GAGGGAACGAGAGATTAAGATCAGACAGCAACGATACCCAGTTTGTCTGCCATGTCTGGC 636
DB 738 GATCAGAGACAGACAGACAGACAGACAGTACAGTACCACTGGAGCCTTTCGCATGATGTC 797
QY 637 GCGCTCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 696
DB 798 ACCCTCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857
QY 697 GAAAGAGAGATGCGGCTTGAACACCTATGCTATGCTATGCTATGCTATGCTATGCTATG 756
DB 858 GCCCAGACACCGAGACCGCTCAATTAATCTATCTCTTGTGGAGCGAGTGTGCTAACCA 917
QY 757 ATCTGAGGCTCATCTTGTGGCTCACCCCAAGGAGAGATCAGCAAGATATGATGACAGT 816
DB 918 GTGACAGATTCAGATATGTACAAAGAGGCAAGTGTGATGATGATGATGATGATGATGAT 977
QY 817 GCGGTGTGTGAGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 876
DB 978 GCGAGGCTGTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1037
QY 877 CTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936
DB 1038 GCGGCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1097

```

```

QY 937 CCGGCGTGTGTGAACGAGTGTGGGATTTCCACAGCTCATATGAGGCTACAACTTC 996
DB 1098 ACCGATCTCTGAGAGTCCCGCTTGCATTCAGAGACATGTGGCATTCACAACTGCAC 1157
QY 997 AGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1035
DB 1158 GGCATTCGTGGATCATATAGCGGCAATGTGGTGTGTGTGTGTGTGTGTGTGTGTGT 1196

```

```

RESULT 6
US-10-027-632-282039

```

```

; Sequence 282039, Application US/10027632
; GENERAL INFORMATION:

```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

```

```

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

```

```

; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

```

```

; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20

```

```

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29

```

```

; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24

```

```

; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23

```

```

; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

```

```

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09

```

```

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 282039
; LENGTH: 505

```

```

; TYPE: DNA
; ORGANISM: Human

```

```

US-10-027-632-282039

```

```

Query Match      11.5%; Score 144.6; DB 13; Length 505;
Best Local Similarity 86.9%; Pred. No. 7.1e-36;
Matches 159; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```

```

QY 758 TCTCAGGCTCATCTTGGCTCACCCCAAGGAGATCAGCAAGATTAATGTCACAGTGC 817
DB 241 TCTCTCTACTCTTCTTCTTACCCACAGCTATTTCTTGCAGACTTATGTGCACAGTGC 300

```

```

QY 818 CGGTGTGTGACAGAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 877
DB 301 CGGTGTGTGACAGAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360

```

```

QY 878 TTGCTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 937
DB 361 TTGCTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420

```

```

QY 938 GGG 940
DB 421 GGG 423

```

```

RESULT 7
US-10-027-632-282040

```

```

; Sequence 282040, Application US/10027632
; GENERAL INFORMATION:

```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

```

```

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

```





```

: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecmics-x-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 18306
: LENGTH: 123
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL031284.9
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 2.5
: OTHER INFORMATION: NT HIT: g11139304, EVALUE 2.00e-63
: OTHER INFORMATION: SWISSPROT HIT: P18577, EVALUE 2.00e-20
: OTHER INFORMATION: EST_HUMAN HIT: T84327.1, EVALUE 3.00e-53
: US-09-864-761-18306

```

OY	487	ACAGCTCACCACATTAATATATGACATCTACGTTTCGCGGCTATTTTGGGCT	546
Db	123	ACAGCTACCAATATGAACTTACAGGCACTTTTACGTTTCGCGGCTATTTTGGGCTACT	64
OY	547	GTGGCTGTGCTGTCCCAAGGCTCTTACCAGGAGGAGGAGATTAAGATCAGACGA	606
Db	63	GTGGCTGTGCTGTCCCAAGGCTCTTACCAGGAGGAGGAGATTAAGATCAGACGA	4
OY	607	ACG 609	
Db	3	ACG 1	
RESULT 13			
US-10-012-542-75			
; Sequence 75, Application US/10012542			
; Publication No. US20030044851A1			
; GENERAL INFORMATION:			
; APPLICANT: Ruben et al.			
; TITLE OF INVENTION: 94 Human Secreted Proteins			
; FILE REFERENCE: P7029p1			
; CURRENT APPLICATION NUMBER: US/10/012.542			
; PRIOR FILING DATE: 2001-12-12			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325			
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507			
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508			
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509			
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510			
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112			
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113			
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22			
; NUMBER OF SEQ ID NOS: 532			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 75			
; LENGTH: 1650			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-012-542-75			
Query Match 8.7%; Score 109.6; DB 14; Length 1650;			
Best Local Similarity 48.7%; Pred. No. 2.7e-24;			
Matches 327; Conservative 0; Mismatches 344; Indels 1; Gaps 1			
OY	364	TGGGTCGATCTCGATGGATGCTGCTCTTTGGGGAAGGTAACCTTGGCGCATTTGGGGG	423
Db	33	TCTGTCGCGTGCGCTTTGGGGCAATTTGGGTAAGTAGGCCCACTTACGCTGCTCATC	92
OY	424	ATGTGCTGTGGAGGTGACAGACTTTAAGCAACTGAGGATGTCATCATGATATCTTC	483
Db	93	ATGACTTTCTTCCAAAGTGACCTCTTTCGGTGTGAATGAATCATCTCTCTTAACCTGCTA	152
OY	484	AACACAGACTACACATATGATATGACATCTACGTTTGGCAGCCTATTTTGGGCTG	543
Db	153	AAGGTGAAGATGACGAGGAGCTCCATGACATCCACATTTTGGGCGCTACTTTGGGCTC	212
OY	544	TCTGTGGCTGTGGCTGCTGCGCAAAAGGCTCTACCCGAGGAGGAGGAGGATTAAGATCAGACA	603
Db	213	ACAGTGACCCGGATCTCTTACCGAGCAACCTTAGACAGACGAGAGGAGACAGAAATTTT	272
OY	604	GCAACGATACCCAGTTTGTGTGACATGCTGGGCGGCGCTCTTCTGTGATGTCTTGCGCA	663
Db	273	GTGTACCAAGTGGAGACTCTTTCCATGATTTGGCACCCCTCTCTGATGTACTGGGCC	332
OY	664	AGTTTCACACTCTGCTCTGCTGTGAAGTCCATTCGAAAGAAATATCCGTTGTTCAACAC	723
Db	333	AGCTTCAACTCAGCCATATCTTACCATGGGAGCACCCAGACCGGCGCATCAACACC	392











QY	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
481	TTCAACAGACTACACATATGATGACATTCACGTTTCGACGCTATTTCGGG				
508	TTCAACAGACTACACATATGATGACATTCACGTTTCGACGCTATTTCGGG				
541	CTGTCTGTGGCTGTGCTGTGCTGCAAGCTCTTACCCAGGAGGAGATTAATGATCAG				
568	CTGACTGTGGCTGTGCTGTGCTGCAAGCTCTTACCCAGGAGGAGATTAATGATCAG				
601	ACAGCAACGATACCCAGTTTGTCTGCGCATGCTGGGCGCCCTCTTCTTGT				
628	AGACCAACGATACCCAGTTTGTCTGCGCATGCTGGGCGCCCTCTTCTTGT				
RESULT 4	B0655978		685 bp	linear	EST 30-SEP-2002
LOCUS	cl04g11.21		Hembase; Erythroid Precursor Cells (LCB:cl library)		Homo sapiens cDNA clone cl04g11 5', mRNA sequence.
ACCESSION	B0655978				
VERSION	B0655978.1		GI:23368160		
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 685)				
AUTHORS	Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.				
TITLE	Gene Expression in Human Erythroid Precursor Cells				
JOURNAL	Unpublished				
COMMENT	Contact: Jeffery L. Miller Laboratory of Chemical Biology National Institute of Diabetes and Digestive and Kidney Diseases Building 10, Room 9B17, National Institutes of Health, Bethesda, MD 20892, USA Tel: 301 402 2373 Fax: 301 435 5148 Email: jml7fe@nih.gov The cl library was constructed by Alexander Gubin, Ph.D. in the Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or analyses by National Institutes of Health Intramural Sequencing Center (NISC). More information available at: http://hembase.nidk.nih.gov Plate: 04 row: g column: 11 Seq primer: 5' lambda-Triplex2 Sequencing Primer. Location/Qualifiers 1. 685 /organism="Homo sapiens" /mol_type="RNA" /db_xref="taxon:9606" /clone="cl04g11" /sex="unknown" /tissue_type="blood" /cell_type="Erythroid Precursor Cells" /cell_line="Primary Culture of Peripheral Blood Mononuclear Cells" /dev_stage="Precursor erythroblasts; GPA++" /lab_host="DH5alpha" /clone_id="Hembase; Erythroid Precursor Cells (LCB:cl library)" /note="Organ: blood; Vector: pTriplex2; Site_1: SfiI; Site_2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos (N-terminal )-biotin-CTC-CAC-CCG-AAG-CTT-G-(C-terminal) and (N-terminal)-biotin-CT/G-C/T-GAA-GTT-AGG-A-(C-terminal ). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH				

Intramural Sequencing Center (NISC):									
Http://www.nisc.nih.gov/									
BASE COUNT 129 a 183 c 200 g 173 t									
ORIGIN									
Query Match	49.2%	Score 617.2	DB 13	Length 685					
Best Local Similarity	98.0%	Pred. No. 4.8e-151							
Matches 625	Conservative 0	Mismatches 13	Indels 0	Gaps 0					
QY	1	ATGAGCTCTAAGTACCCGCGCTGTCTGTCCGGCGCTGCCTCCCTCTCTGGGCCCTTAACACTG	60						
Db	48	ATGAGCTCTAAGTACCCGCGCTGTCTGTCCGGCGCTGCCTCCCTCTCTGGGCCCTTAACACTG	107						
QY	61	GAAGCAGCTCATATTTCTCTCTTTATTTTTTAAACCCTATGACGCTTCTTAGAGAT	120						
Db	108	GAACAGCTCATATTTCTCTCTTTATTTTTTAAACCCTATGACGCTTCTTAGAGAT	167						
QY	121	CAAAAGGGGCTCGTGGGCAATCTATCAAGTTGGCCAAAGATCTGACCCTGATGGGGCCATT	180						
Db	168	CAAAAGGGGCTCGTGGGCAATCTATCAAGTTGGCCAAAGATCTGACCCTGATGGGGCCATT	227						
QY	181	GGCTTGGGCTTCTCTACCTCGAGTTTCCGGAGACACACTGGAGCAGTGGGCTTCAAC	240						
Db	228	GGCTTGGGCTTCTCTACCTCGAGTTTCCGGAGACACACTGGAGCAGTGGGCTTCAAC	287						
QY	241	CTCTTCATGCTGGGCTTGTGTGTCAGTGGGCAATCTCTGGAAGGCTTCTGAGCCAG	300						
Db	288	CTCTTCATGCTGGGCTTGTGTGTCAGTGGGCAATCTCTGGAAGGCTTCTGAGCCAG	347						
QY	301	TTCCCTTCTGGGAAGGTGTATATACACTGTTCAGTATTTGGGCTGGCCACCATGATGCT	360						
Db	348	TTCCCTTCTGGGAAGGTGTATATACACTGTTCAGTATTTGGGCTGGCCACCATGATGCT	407						
QY	361	TTGTGCGTGCATCTCAGTAGTATGCTGTCTGGGGAAGTCAACTTGGGCGACATTTGGTG	420						
Db	408	GTGTGCGTGCATCTCAGTAGTATGCTGTCTGGGGAAGTCAACTTGGGCGACATTTGGTG	467						
QY	421	GTGATGTGCTGTGTGAGGTGACAGCTTTAGGCAACCTGAGGATGTCATGATATATC	480						
Db	468	GTGATGTGCTGTGTGAGGTGACAGCTTTAGGCAACCTGAGGATGTCATGATATATC	527						
QY	481	TTCAACACAGCTACCATATGATATGACATCTACGTTTCGACGCTATATTTGGG	540						
Db	528	TTCAACACAGCTACCATATGATATGACATCTACGTTTCGACGCTATATTTGGG	587						
QY	541	CTGACTGTGGCGCTGCTGCCTCCAAAGCCTTACCCGAGGAGGAGGATTAAGATCAG	600						
Db	588	CTGACTGTGGCGCTGCTGCCTCCAAAGCCTTACCCGAGGAGGAGGATTAAGATCAG	647						
QY	601	ACAGCAACGATATCCAGTTTGTCTGCCATGCTGGGGCG	638						
Db	648	AGAGCAACGATATCCAGTTTGTCTGCCATGCTGGGGCG	685						
RESULT 5									
BX406634/c 1000 bp mRNA linear EST 13-MAY-2003									
LOCUS BX406634 Homo sapiens FETAL LIVER Homo sapiens cDNA clone									
DEFINITION CSDDM013YA17 3-PRIME, mRNA sequence.									
ACCESSION BX406634									
VERSION BX406634.1 GI:30656378									
SOURCE EST.									
ORGANISM Homo sapiens (human)									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE 1 (bases 1 to 1000)									
AUTHORS Li,W.B., Gruber,C., Jessee,J., and Polayes,D.									
JOURNAL Full-length cDNA libraries and normalization									
COMMENT Unpublished									
Contact: Genoscope									
Genoscope - Centre National de Sequencage									
BP 19191006 EVRY cedex - France									

TCACCTACATTGTGCTGCTGCTTCATACTGCTGGAACGGCAATGGCATGA

Wynshaw-Borja A   Yoshida K   Hasegawa Y   Kawai H   Kohsunkul S  
Toyo-oka, K., Wang, K.H., Weltz, C., Whitaker, C., Wilming, L.,

FEATURES	source	COMMENT	TITLE	JOURNAL	REFERENCE	AUTHORS
CDS	1. 1457 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:9630028A03" /db_xref="taxon:10090" /clone="9630028A03" /tissue_type="cerebellum" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="16 days neonate" 10. 1266 /note="unnamed protein product; Rhesus blood group CE and D (MGDI)MG1:1202882, GB NM_011270, evidence: BLASTN, 99%, match=1257) putative" /codon_start=1 /protein_id="BAC37610.1" /db_xref="GI:26347923" /translation="MGSKYRPSLRCLPLALDELQAFSLFCFFIPHDVAQVDHREPMASVQYLNLTLMALGFGFLSSFRSRSSWVAANLEMLAGVGQITLDFHQLQVLAQWKNLNINLSIDIOIAMSTLPVLISAGVALKYNLVQIVYVIMLEAMAGCAIFAEDEKVFMTKHIIIMGHVFGAYGLTVAMWLSRSRLPRVGENAOFKRYOMARSSSLFALNGITVLFIETWPAINSALEIGTRKRNRAVENGTALVALAVATATMSKALSHPOGKTIMVTHINNAVLACGAVAGACCLISSPWSWVGLTAGLISITGAKCPRCLNHNMLONSSGHIYITGGLPLGLALTYICIDIVTEPKSSDMLITQYTVTHIGALSFVAMGMVGLTLCILLSVRVMPRAKRYFDDTETWEPFLAVGF"	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> URL: <a href="http://fantom.gsc.riken.go.jp/">http://fantom.gsc.riken.go.jp/</a> . Location/Qualifiers	Functional annotation of a full-length mouse cDNA collection Nature 405 (6821), 685-690 (2001)	Nature 405 (6821), 685-690 (2001)	Nature 405 (6821), 685-690 (2001)	Hayashizaki, Y. and Hayashizaki, Y. 21085860 PUBMED 11217851
polyA_signal	1457 /note="putative"		The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1457) Fukuda, S., Alizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Hayashida, K., Hayatsu, N., Hiyamoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotoh, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawaji, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.			
polyA_site	1457 /note="putative"		Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> , Tel: 81-45-503-9216, Fax: 81-45-503-9216)			
polyA_site	1457 /note="putative"		Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> , Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
polyA_site	1457 /note="putative"		CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> URL: <a href="http://fantom.gsc.riken.go.jp/">http://fantom.gsc.riken.go.jp/</a> . Location/Qualifiers			
polyA_site	1457 /note="putative"		Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> , Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
polyA_site	1457 /note="putative"		CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> URL: <a href="http://fantom.gsc.riken.go.jp/">http://fantom.gsc.riken.go.jp/</a> . Location/Qualifiers			
polyA_site	1457 /note="putative"		Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> , Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
polyA_site	1457 /note="putative"		CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> URL: <a href="http://fantom.gsc.riken.go.jp/">http://fantom.gsc.riken.go.jp/</a> . Location/Qualifiers			
polyA_site	1457 /note="putative"		Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> , Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
polyA_site	1457 /note="putative"		CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> URL: <a href="http://fantom.gsc.riken.go.jp/">http://fantom.gsc.riken.go.jp/</a> . Location/Qualifiers			
polyA_site	1457 /note="putative"		Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> , Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
polyA_site	1457 /note="putative"		CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> URL: <a href="http://fantom.gsc.riken.go.jp/">http://fantom.gsc.riken.go.jp/</a> . Location/Qualifiers			
polyA_site	1457 /note="putative"		Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> , Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
polyA_site	1457 /note="putative"		CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> URL: <a href="http://fantom.gsc.riken.go.jp/">http://fantom.gsc.riken.go.jp/</a> . Location/Qualifiers			
polyA_site	1457 /note="putative"		Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> , Tel: 81-45-503-9222, Fax: 81-45-503-921			

Query Match	48.0%	Score 602.2	DB 11	Length 1457
Best Local Similarity	69.7%	Pred. No.5.6e-147		
Matches	884	Conservative	0	Mismatches 358; Indels 27; Gaps 4;
QY	1	ATGAGCTCTAAGTACCCGGGCTGTGTGCGGGCTGCTGCCCTTCTGGGACCTTAACATG	60	
Db	10	ATGGGCTTAAGTACCCAGGCTCCCTCGGCTGTGCTCCGCTTGGGCTTGGAGCTA	69	
QY	61	GAAGACGCTCATCTTCTCTCTCTATTTTTTAAACCATAAGACGCTTCTTAGAGAT	120	
Db	70	CAGACAGCTTTAAGTCTCTCTCTGTTTTTTCATCCCCCAGCACAGCCAGGTGAT	129	
QY	121	CAAAAGGGGCGTGGGCAATCTATCAATTGGCCAAAGATCTAGACGTATGAGCCCAATT	180	
Db	130	CACA---GGTTCATAGGCGAGCTATCAAGTCTCCGGAAATTGACCTTATGCAACCTTG	186	
QY	181	GCGTTGGGCTTCTCTACCTCGATGTTTCCGAGACACACAGCTGGAGCATGTGGCTTCAAC	240	
Db	187	GGCTTCGGCTTCTCTGTCTCTGCTCTTGGAGACACACAGCTGGAGCATGTGGCTTCAAC	246	
QY	241	CTTTTCATGCTGGGCTTGTGTGAGTGGCAATCTGTGTGAGCGCTTCTTGAAGCAG	300	
Db	247	CTTTTCATGTTGGCCTTCGGGGGTGAGGGAACAATCTTGTGACCAATTTCTGTGGCCAG	306	
QY	301	TTCCCTTGTGGGAAGGTGATCAACAATGTTTCAGTATGTGGGCGGACCATAGTGTCT	360	
Db	307	GTCTCCCATGGAACAAGATCAACATGTGTTCACATCCAGATACAGTACATGAGCAC	366	
QY	361	TTGTGCGTGCATCTCAGTGGATGCTGTCTTGGGGAAGTCACTTGTGGCAGTGTGTG	420	
Db	367	TTACTGTGTGATCTCAGGGGGGCTGTCTCGGGGAAGGTCAACTGTGTGACGTGACC	426	
QY	421	GTATGTGTGTGTGAGGTACACGCTTTAGGCAACTGAGGATGTATCATGTAATATC	480	
Db	427	GTGATGTGTGTGTGAGGAGGCAATGAGCTTTGTGTGCATCAATTTGGCCGAGAGAGGTC	486	
QY	481	TTCAACAACAGCTACCAATGAATGATGTGACATCTCTCGTGTGGCACTAATTTTGG	540	
Db	487	TTCAAAATGACAGAACACATCATGTGATGACAGGGCACGTGTTGGGGCTATTTTGGG	546	
QY	541	CTGTCTGTGGCCTGTGCTCCCAAGGCTTACCCGAGGAGACGAGGATTAAGATCAG	600	
Db	547	CTAATGTGTGGCTGTGCTTCCAGATCTGTGCCACAGAGATGTGGTGAAGACGCCAG	606	
QY	601	ACA-----GCAACGATACCCAGTTGTGTGCTCATGTGTGGGCGCTCTTTC	645	
Db	607	ACAGAAGAGTTCAATGAGTACGAGCTCAGTCTGTTTGGCATGTGGGACCTCTTCT	666	
QY	646	TTGTGATGTCTTGGCCCAAGTTTCAACTCTCTCTGTGGGAAGTCAATGGAAGAAAG	705	
Db	667	TTGTGATGTCTTGGCCCAAGTATCAACTCTCTCTCTGTGGGAAGGACAAAGAAAAAG	723	
QY	706	AATGCCGTGTCAACACCTACTATGTGTGATGACAGTCAAGCGTGTGACAGCCATCTCAGG	765	
Db	724	AATGCTGTGTCAACACCTACTATGAGCCCTGTGACAGTCAAGCGTGTGACAGCCATCTCAGG	783	
QY	766	TCATTCCTTGGCTCACCCCAAGGGAAGATCAAGCAAGCTTATGTGTACAGTGTGGGTGTG	825	
Db	784	TCACCCCTGATCACCCCTCAAGGGAAGTCAACATGTGTCAATCACAAGTCAAGTGTG	843	
QY	826	GAGAGAGGCGGCTGTGGTACCTGCGTACCTGATCCCTTTCGCTGTGGCTTGCATG	885	
Db	844	GCAAGGAGGCGGCTGTGGTACCTGCGTACCTGATCCCTTTCGCTGTGGCTTGCATG	903	
QY	886	GTCCTGTGTCTTGTGGCTGTGATCTCCGTGGGGAGGCCAAGTACCTGCGGGGTGT	945	
Db	904	GTGCTGTGGGCTTACATAGCTGGGTGTATCTCCATCTGTGGGAGCCAAAGTGTCCACGGGGTGT	963	
QY	946	TGTAAACGAGTGTGGGATTTCCCAACAGCTTCATCATGGGCTTCAACTTACGTTGTGTG	1005	
Db	964	TTGAAACCAATGTGTGAGAACTTCAGATGG-----GATCCACTACACTTGTGGGTGGCG	1017	



	BASE COUNT	
ORIGIN	136 a 186 c 209 g 175 t	
Query Match	47.6%; Score 597.4; DB 13;	Length 706;
Best Local Similarity	97.4%; Pred. No. 7,7e-146;	
Matches 607; Conservative	0; Mismatches 16; Indels 0; Gaps 0;	
OY	1 ATGAGCGCTTAAGTACCCGGGCTCMTCCGGGGGTGCCTGGCCCCCTTGSGGCCCTAAACTG 60   Db	84 ATGAGCGCTTAAGTACC GGCGGTCTGTCGGCGCTGCTGCCCCCTTCTGGGCCCTTAACCTG 143
OY	61 GAAGCACGCTCATTTCTCTCTATTTTTTACCACATATGACGCTTCCCTTAGAGGAT 120   Db	144 GAAGCACGCTCATATTCTCTCTATTATTTTTTACCACATATGACGCTTCCCTTAGAGGAT 203
OY	121 CAAAAGGGGCTCGTGGAATCCCTATCAAGTTGGCCAGAATCTGACCGTGATGGCGCCATT 180   Db	204 CAAAAGGGGCTCGTGGAATCCCTATCAAGTTGGCCAGAATCTGACCGTGATGGCGCCCTT 263
OY	181 GGGCTTGGGGCTCCGACACTCGAGTTTCCGGAGACAGACGCTGGAGCACTGGTGGCCCTTAAC 240   Db	264 GGGCTTGGGGCTCCGACACTCAAATTTCCGGAGACAGCTGGAGCACTGGTGGCCCTTAAC 323
OY	241 CTCTTCATGCTGGGCGCTTGGTGTGCACTGGGCAATCCTGCTGACGCGCTTCTTAGACCAAG 300   Db	324 CTCTTCATGCTGGGCGCTTGGTGTGCACTGGGCAATCCTGCTGACGCGCTTCTTAGACCAAG 383
OY	301 TTCCCCTTGGGAAGGTGTGATCATCACTGTTCAATAATTCGGCTGGCCACCATGAGTGCT 360   Db	384 TTCCTCTCGTGGGAAGGTGTGATCATCACTGTTCAATAATTCGGCTGGCCACCATGAGTGCT 443
OY	361 TTTGCGGTGCGATCAGTGGATGAGTCGTCTGGGGGAAGGCACTTGGGCGAGTTGGT 420   Db	444 ATGTGCGGTGATCTAGCGGGGTGCTCTTGGGAAGGCACTTGGGCGAGTTGGT 503
OY	421 GTGATGTGCTGTGTGAGAGTGACAGCTTTAAGGCAACCTGAGAGATGTCATCAATAATATC 480   Db	504 GTGATGTGCTGTGTGAGAGTGACAGCTTTAAGGCAACCTGAGAGATGTCATCAATAATATC 563
OY	481 TTCAACAACAGACTAACATGAAATGATGACATCTAAGCTGTTGCGAGCCTATTTTGGG 540   Db	564 TTCAACAACAGACTAACATGAAATGATGACATCTAAGCTGTTGCGAGCCTATTTTGGG 623
OY	541 CTGCTGTGGGCTGTGTGCTGCCCAAAGCCTTACCAGAGGAAGGAGAGATAAAGATCAG 600   Db	624 CTGCTGTGGGCTGTGTGCTGCCCAAAGCCTTACCAGAGGAAGGAGAGATAAATGATCAG 683
OY	601 ACAGCAACGATACCAAGTTTGTGCT 623   Db	684 AGAGCAACGATACCAAGTTTGTGCT 706
RESULT 8		
B0662207		
LOCUS	631 bp mRNA Linear EST 30-SEP-2002	
DEFINITION	c182h04.z1 Hembase; Erythroid Precursor Cells (LCB:c1 library) Homo sapiens cDNA clone c182h04 5' , mRNA sequence.	
ACCESSION	B0662207	
VERSION	B0662207.1 GI:23374389	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

REFERENCE 1 (bases 1 to 631)  
 AUTHORS Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.  
 TITLE Gene Expression in Human Erythroid Precursor Cells  
 JOURNAL Unpublished  
 COMMENT Contact: Jeffery L. Miller  
 Laboratory of Chemical Biology  
 National Institute of Diabetes and Digestive and Kidney Diseases  
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
 20892, USA  
 Tel: 301 402 2373  
 Fax: 301 435 5148  
 Email: jml7@nih.gov  
 The 'cl' library was constructed by Alexander Gubin, Ph.D. in the  
 Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or  
 analyses by National Institutes of Health Intramural Sequencing  
 Center (NISC). More information available at:  
<http://hembase.nidk.nih.gov>  
 Plate: 82 row: h column: 04  
 Seq primer: 5' lambda-TripLex2 Sequencing primer.  
 Location/Qualifiers  
 1..631  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="cl82h04"  
 /sex="unknown"  
 /tissue\_type="blood"  
 /cell\_line="Erythroid Precursor Cells"  
 /cell\_line="Primary Culture of Peripheral Blood Mononuclear Cells"  
 /dev\_stage="Precursor erythroblasts; GPA++"  
 /lab\_host="DH5alpha"  
 /clone\_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"  
 /note="Organ: blood; Vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiI. A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos (N-terminal )-biotin-6TC-CAC-CCG-ANG-CTT-G-(C-terminal) and (N-terminal)-biotin-CTT-C/T-GAA-GTT-CTC-AGG-A-(C-terminal) . Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC).  
<http://www.nisc.nih.gov/>."

BASE COUNT 114 a 163 c 188 g 165 t 1 others

ORIGIN

Query Match 44.7%; Score 560.4; DB 13; Length 631;  
 Best Local Similarity 99.6%; Pred. No 3.9e-136;  
 Matches 561; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATAGGCTTAAGTACCGCGGCTGTCCGGCGCTGCTGCTGCTGCGGCTTAACACTG 60  
 |||||||  
 Db 69 ATAGGCTTAAGTACCGCGGCTGTCCGGCGCTGCTGCTGCGGCTTAACACTG 128  
 |||||||  
 QY 61 GAAGCAGCTCTATTCT 120  
 |||||||  
 Db 129 GAAGCAGCTCTATTCT 188  
 |||||||  
 QY 121 CAAGAAGGGCTGCTGATCTCTATCAAGTTGGCAAGATCTGACCGGTGAGGGGCATT 180  
 |||||||  
 Db 189 CAAGAAGGGCTGCTGATCTCTATCAAGTTGGCAAGATCTGACCGGTGAGGGGCATT 248  
 |||||||  
 QY 181 GGGTTGGGCTTCTCAGCTCGAGTTTCCGAGACAGCTGAGCAGTGTGGCTTCAAC 240  
 |||||||  
 Db 249 GGGTTGGGCTTCTCAGCTCGAGTTTCCGAGACAGCTGAGCAGTGTGGCTTCAAC 308

QY 241 CTCTTCATGCTGGCGCTTGGTGTGACATTCCTGCTGACGCGCTTCTGAGCCAG 300  
 |||||||  
 Db 309 CTCTTCATGCTGGCGCTTGGTGTGACATTCCTGCTGACGCGCTTCTGAGCCAG 368  
 |||||||  
 QY 301 TTCCCTTCTGGGAAGGTGCTATCAGCTGTTCAGTATTTGGCTGGCCACCATGAGTCT 360  
 |||||||  
 Db 369 TTCCCTTCTGGGAAGGTGCTATCAGCTGTTCAGTATTTGGCTGGCCACCATGAGTCT 428  
 |||||||  
 QY 361 TTTCGGGTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
 |||||||  
 Db 429 TTTCGGGTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 488  
 |||||||  
 QY 421 GTGATGCTGCTGTGAGTGTGACAGCTTTAGCCAACTGAGATGATCAGTAATATC 480  
 |||||||  
 Db 489 GTGATGCTGCTGTGAGTGTGACAGCTTTAGCCAACTGAGATGATCAGTAATATC 548  
 |||||||  
 QY 481 TTCAACACAGACTACACATGACATGATGATGATGATGATGATGATGATGATGATG 540  
 |||||||  
 Db 549 TTCAACACAGACTACACATGACATGATGATGATGATGATGATGATGATGATGATG 608  
 |||||||  
 QY 541 CTGCTGTGGCTGCTGCTGCTGCC 563  
 |||||||  
 Db 609 CTGCTGTGGCTGCTGCTGCC 631  
 |||||||

RESULT 9  
 B0665450 632 bp mRNA linear EST 30-SEP-2002  
 LOCUS cl131d04.z1 Hembase; Erythroid Precursor Cells (LCB:cl library)  
 DEFINITION Homo sapiens cDNA clone cl131d04 5', mRNA sequence.  
 ACCESSION B0665450  
 VERSION B0665450.1 GI:23377637  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 632)  
 Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.  
 TITLE Gene Expression in Human Erythroid Precursor Cells  
 JOURNAL Unpublished  
 COMMENT Contact: Jeffery L. Miller  
 Laboratory of Chemical Biology  
 National Institute of Diabetes and Digestive and Kidney Diseases  
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
 20892, USA  
 Tel: 301 402 2373  
 Fax: 301 435 5148  
 Email: jml7@nih.gov  
 The 'cl' library was constructed by Alexander Gubin, Ph.D. in the  
 Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or  
 analyses by National Institutes of Health Intramural Sequencing  
 Center (NISC). More information available at:  
<http://hembase.nidk.nih.gov>  
 Plate: 131 row: d column: 04  
 Seq primer: 5' lambda-TripLex2 Sequencing primer.  
 Location/Qualifiers  
 1..632  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="cl131d04"  
 /sex="unknown"  
 /tissue\_type="blood"  
 /cell\_line="Erythroid Precursor Cells"  
 /cell\_line="Primary Culture of Peripheral Blood Mononuclear Cells"  
 /dev\_stage="Precursor erythroblasts; GPA++"  
 /lab\_host="DH5alpha"  
 /clone\_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"  
 /note="Organ: blood; Vector: pTriplex2; Site\_1: SfiI;

Site\_2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA library construction kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos (N-terminal (N-terminal)-biotin-CAC-CCG-AGG-CTT-G-(C-terminal) and (N-terminal)-biotin-C(T/C)-T-GAA-GTT-CTC-AGG-A-(C-terminal)). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC); <http://www.nisc.nih.gov/>.

BASE COUNT 115 a 171 c 185 g 161 t  
ORIGIN

Query Match 44.6%; Score 559; DB 13; Length 632;  
Best Local Similarity 98.3%; Pred. No. 9e-136;  
Matches 565; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAATACCCGGGCTGTCCGGGCTGCTGCGCCCTGAGCCCTTAACATCG 60
    |||||||
Db 58 ATGAGCTCTAATACCCGGGCTGTCCGGGCTGCTGCGCCCTGAGCCCTTAACATCG 117
QY 61 GAAGAGCTCTATTCCTCTCTTATTTTATTTTACCACATATAGAGCTCTTAGAGAT 120
    |||||||
Db 118 GAAGAGCTCTATTCCTCTCTTATTTTATTTTACCACATATAGAGCTCTTAGAGAT 177
QY 121 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGATCTAGCCGTATGGCGGCATT 180
    |||||||
Db 178 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGATCTAGCCGTATGGCGGCATT 237
QY 181 GGCTTGGGCTTCCCTACCTCGAGTTTCCGGAGACACAGCTGAGCAGTGGCTTCAAC 240
    |||||||
Db 238 GGCTTGGGCTTCCCTACCTCGAGTTTCCGGAGACACAGCTGAGCAGTGGCTTCAAC 297
QY 241 CTTTCATGCTGGGCTGTGGTGTGCAAGTGGCAATCTGAGAGGCTTCCGAGACCG 300
    |||||||
Db 298 CTTTCATGCTGGGCTGTGGTGTGCAAGTGGCAATCTGAGAGGCTTCCGAGACCG 357
QY 301 TTCCCTTCTGGGAGGCTGTCATCACTGTTCAGTATTCGGCTGGCCACATGAGTCT 360
    |||||||
Db 358 TTCCCTTCTGGGAGGCTGTCATCACTGTTCAGTATTCGGCTGGCCACATGAGTCT 417
QY 361 TTGTGGGTGCTATCTCACTGATGCTGTCTTGGGAGGTCACCTGGCGAGTTGGT 420
    |||||||
Db 418 AGCTGTGCTGCTATCTCACTGATGCTGTCTTGGGAGGTCACCTGGCGAGTTGGT 477
QY 421 GTGATGGTGTGGTGGAGGTGACAGCTTAGGCAACCTGAGATGTCTACGTATATTC 480
    |||||||
Db 478 GTGATGGTGTGGTGGAGGTGACAGCTTAGGCAACCTGAGATGTCTACGTATATTC 537
QY 481 TTCAACACAGACTACACATGAACATGATGACATCTAGCTGTGCGACACCTATTTTGGG 540
    |||||||
Db 538 TTCAACACAGACTACACATGAACATGATGACACCTTCTAGCTGTGCGACACCTATTTTGGG 597
QY 541 CTGTCTGTGGCTGTGGCTGCGCAAGCTCTTACC 575
    |||
Db 598 CTGTCTGTGGCTGTGGCTGCGCAAGCTCTTACC 632

```

RESULT 10

LOCUS BU664659 613 bp mRNA linear EST 30-SEP-2002  
DEFINITION Homo sapiens Erythroid Precursor Cells (LCB:cl library)  
ACCESSION BU664659  
VERSION BU664659.1 GI:23376845  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L.  
TITLE Gene Expression in Human Erythroid Precursor Cells  
JOURNAL Unpublished  
COMMENT Contact: Jeffery L. Miller  
Laboratory of Chemical Biology  
National Institute of Diabetes and Digestive and Kidney Diseases  
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
20892, USA  
Tel: 301 402 2373  
Fax: 301 435 5148  
Email: [jm@leish.gov](mailto:jm@leish.gov)  
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or analyses by National Institutes of Health Intramural Sequencing Center (NISC). More information available at: <http://hembase.nidk.nih.gov>  
Plate: 120 row: e column: 10  
Seq primer: 5' lambda-Triplex2 Sequencing Primer.

#### FEATURES

##### source

1. 613  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="cl120e10"  
/sex="unknown"

/tissue\_type="blood"  
/cell\_type="Erythroid Precursor Cells"  
/cell\_line="Primary Culture of Peripheral Blood Mononuclear Cells"  
/dev\_stage="Precursor erythroblasts; GPA++"  
/lab\_host="DH5alpha"  
/clone\_11b="Hembase; Erythroid Precursor Cells (LCB:cl library)"

/note="Organ: blood; Vector: pTriplex2; site\_1: SfiI; site\_2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA library construction kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos (N-terminal (N-terminal)-biotin-CAC-CCG-AGG-CTT-G-(C-terminal) and (N-terminal)-biotin-C(T/C)-T-GAA-GTT-CTC-AGG-A-(C-terminal)). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC); <http://www.nisc.nih.gov/>."

BASE COUNT 110 a 159 c 181 g 161 t 2 others  
ORIGIN

Query Match 44.0%; Score 552; DB 13; Length 613;  
Best Local Similarity 99.6%; Pred. No. 6.1e-134;  
Matches 552; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAGCTCTAATACCCGGGCTGTCCGGGCTGCTGCGCCCTGAGCCCTTAACATCG 60
    |||||||
Db 60 ATGAGCTCTAATACCCGGGCTGTCCGGGCTGCTGCGCCCTGAGCCCTTAACATCG 119
QY 61 GAAGAGCTCTATTCCTCTCTTATTTTATTTTACCACATATAGAGCTCTTAGAGAT 120
    |||||||
Db 120 GAAGAGCTCTATTCCTCTCTTATTTTATTTTACCACATATAGAGCTCTTAGAGAT 179
QY 121 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGATCTAGCCGTATGGCGGCATT 180
    |||||||
Db 180 CAAAAGGGGCTGTGGCATCTATCAAGTTGGCCAAAGATCTAGCCGTATGGCGGCATT 239
QY 181 GGCTTGGGCTTCCCTACCTCGAGTTTCCGGAGACACAGCTGAGCAGTGTGGCTTCAAC 240
    |||||||

```

Db	240	GGCTTGGGCTTCCTCACCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGGCTTCAAC	299
Oy	241	CTCTTCATGCTGGCCCTTGCTGTGCAAGTGGCAATCTCTGTAGAGCGCTTCTTGAGCCAG	300
Db	300	CTCTTCATGCTGGCCCTTGCTGTGTGCAAGTGGCAATCTCTGTAGAGCGCTTCTTGAGCCAG	359
Oy	301	TTCCCTTTGGGAAGTGGTGCATCACAAGTTCAGATTTCGGCTGGCCACCAATGAGTGGCT	360
Db	360	TTCCCTTTGGGAAGTGGTGCATCACAAGTTCAGATTTCGGCTGGCCACCAATGAGTGGCT	419
Oy	361	TTTGTCGCTCATGATCTCAGTGGATCTGTCTTGGGGAAGTGAACCTTGGACGAGTTGGTG	420
Db	420	TTGTGGGTGCTGATCTCAGTGGATCTGTCTTGGGGAAGTGAACCTTGGACGAGTTGGTG	479
Oy	421	GTGATGTGTGTGGTGGAGGTGACAGCTTTAGGCAACCTTGAGGATGGTTCATCAGTAATATC	480
Db	480	GTGATGTGTGTGGTGGAGGTGACAGCTTTAGGCAACCTTGAGGATGGTTCATCAGTAATATC	539
Oy	481	TTCAACAACAGACTACCAATATGAACTGATGCACATCTACGTGTTGGCAACCTAATTTGGG	540
Db	540	TTCAACAACAGACTACCAATATGAACTGATGCACATCTACGTGTTGGCAACCTAATTTGGG	599
Oy	541	CTGTCTGTGGCCTG 554	
Db	600	CTGTCTGTGGCCTG 613	

RESULT 11	
BU657533	
LOCUS	607 bp mRNA linear EST 30-SEP-2002
DEFINITION	cl25d06.1 Hembase; Erythroid Precursor Cells (LCP:cl library) Homom
DESCRIPTION	sapiens cDNA clone cl25d06 5', mRNA sequence.

SOURCE ORGANISM	Homo sapiens (human)
	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 607)	Gublin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.	Gene Expression in Human Erythroid Precursor Cells

COMMENT	UNPUBLISHED
Contact: Jeffery L. Miller	

Laboratory of Chemical Biology  
National Institute of Diabetes and Digestive and Kidney Diseases  
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
20892, USA  
Tel: 301 402 2373  
Fax: 301 435 5148  
Email: [jm7fe@nih.gov](mailto:jm7fe@nih.gov)  
The 'cl' library was constructed by Alexander Gubin, Ph.D. in the  
Laboratory of Chemical Biology, NIDDK, NIH. DNA sequencing and/or  
analyses by National Institutes of Health Intramural Sequencing  
Center (NISC). More information available at:  
<http://hembase.nidk.nih.gov>  
Plate: 25 row: d column: 06  
Seq primer: 5' lambda-triPLEX2 Sequencing Primer.

FEATURES	Location/Qualifiers
source	1. .607

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c125d06"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood
Nonnucleolar Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_idb="Hembase; Erythroid Precursor Cells (LCB:cl
library)"

```

/note="Organ: blood; Vector: pRIP4X2; Site:1: SfiI; Site:2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1  $\mu$ mol/L peptide nucleic acid (PNA) oligos (N-terminal (PNA)-biotin-GTC-CAC-CGC-AG-C-CTT-G-(C-terminal) and (N-terminal)-biotin-C(CTT)-GGA-GGT-CTC-AGG-A-(C-terminal)). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC; <http://www.nisc.nih.gov/>).

Query Match	43.3%	Score 542.4	DB 13	Length 607
Best Local Similarity	98.0%	Pred. No. 2e-131		
Matches 549	Conservative	0	Mismatches 11	Indels 0
				Gaps 0

QY 1 ATGAGCTCTAAGTACCCGGGCTCTCCGGGCTGCTGGCCCTTGGGGCCCTAACACTG 60  
 Db 48 ATGAGCTCTAAGTACCCGGGCTCTCCGGGCTGCTGGCCCTTGGGGCCCTAACACTG 107

61 GAAGCAGCTCTCATCTCCTCTTCTAATTTTTTTTACCACATATGACGCTTCCTTAGAGCAT 120

Db 108 GAAGCAGCTCTCATTCCTCTCTCAATTTTTTACCACATATGACGCTTCCTTAGAGCAT 167

[illegible]

181 GGCTTGGGCTTCCTCACCTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCCTTCAAC 240

Db 228 GGCATTGGGCTTCCACACTCGAGTTTCCGGAGACACAGCTGGAGCAGTGTGGCCTTCAAC 287

241 CTCCTCATGCTGGCGTTGGTGTGACATGGGCAATCCTGCTGGACGCTTCTTGAGCCAG 300

Db 288 CTCTTCATGCAGGGCGCTTGCTGTGCAGTGGGCATTCGCTGGACGGCTTCCTGAGCCAG 347

[illegible]

361 TTTGCGGCTGCTGATTCCTACGTGGATGCTCTTGGGCGAAAGCTCAACCTTGGGCGGCGAGTTGGCTG 420

Db 408 ATGTCCGCTGCTGATCTCAGCGGCTCTTGGGGAAGTCACACTTGGCCAGTTGGTC 467

421. GTGATGGTCTGGTGGAGGTGACACCTTTAGGCAACCTGAGGATGGTCATCAGTAATATC 480

Db 468 GTATGGTCTGTGGAGGTGACAGCTTTAGGCACCTGAGGATGTCATCAGTAATATC 527

481 TCAACACAGACATACCACATGAACATGATGCACATCTACGTGTCCAGCCATTATTTGGG 540

[illegible]

Db 588 CCGACTGTGGCCCTGGTGCCT 607

AK089642  
LOCUS  
AK089642  
075 bp  
mpDNA  
1 insert  
UMC 05-DEC-200

**DEFINITION** Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone:F830006J05 product:Rhesus blood group CE and D, full

ACCESSION AK089642  
Insert sequence.

**KEYWORDS** HTC; CAP trapper.

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Konno, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Koichiwa, H., Knehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schiraldi, L. M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carlini, P., de Bona, M., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaur, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmberg, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	6 (bases 1 to 975)
PUBMED	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoaka, J., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Konno, S., Konno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sobge, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,

TITLE	Muramatsu, M. and Hayashizaki, Y.
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome-gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers
FEATURES	1..975
source	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="NOD"
	/db_xref="FANTOM:DB:F830006J05"
	/db_xref="taxon:10090"
	/clone="F830006J05"
	/tissue_type="activated spleen"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
CDS	19..>975
	/note="unnamed protein product; Rhesus blood group CE and D (MD) [MGI:1202882, GRI:NM_011270, evidence: BLASTN, 99%, match=1257]
	putative"
	/codon_start=1
	/protein_id="BAC40937.1"
	/db_xref="GI:26354621"
	/translation="MSKTPRSIRCLPLMALELOTAFSLRCFFIPHTDROYDRFM ASYQVRLNTLMAALGFLSSFRHSSVAFNLFMALISVQGITLIDHPLGOVLQ NWKINNLSTOIAITMSTLPLVLSAGAVLKVNIVLTVAIVLMEAFGIRADEKVF KMTETIIVMAGVFAGFGLVAVMVRSLPRVGNQOTEKVQVATSSSLPAMCTL FLIMEPAINSLBETGRKRNVAVENTYRLAVSAVATSMASLSHQGINNWHINA VIAGGAVAPRCLISSPWSIMVGLIADLISIMWAKPRCL"
BASE COUNT	195 a 273 c 271 g 236 t
ORIGIN	
Query Match	39.8%; Score 498.6; DB 11; Length 975;
Best local Similarity	72.4%; Pred. No. 7, 9e-120;
Matches	696; Conservative 0; Mismatches 244; Indels 21; Gaps 3;
OY	1 ATGAGCTCTAAGTACCCGCGGCTGTCGGCGCTGCTGCCCTCTGGGCCCTTAACATG 60
DB	19 ATGGGCTCAAGAACCCACGACGTCCTCCGCTGCTGCTGCGCTGAGCTGA 78
OY	61 GAAGCAGCTCTATCT 120
DB	79 CAGACGCTTTTATGCT 138
OY	121 CAAGAAGGCTCTGTCGATCTATCAAGTTGGCCAGATCTGACCTGATGGCGCCATT 180
DB	139 CACA---GGTTATATGCGAGGATCAAGTCTCCGGAATTTGACCTCATGGAGCTTG 195
OY	181 GCGTTGGCTTCTCTACCTGAGTTTCCGAGACACAGCTGAGCAGTGGCTTCAAC 240
DB	196 GCGTTGGCTTCT 255
OY	241 CTTCTATCTGCGGCTTGTGTCAGTGGGCAATCTGAGGAGGCTTCCGAGCCAG 300
DB	256 CTTCTATCTGCGGCTTGTGTCAGTGGGCAATCTGAGGAGGCTTCCGAGCCAG 315
OY	301 TTCCCTTTCGAGAGGTGCTCATCACATCTTCATTCGGCTGCCACCATGAGTGTCT 360







http://hembase.niddk.nih.gov  
Plate: 49 row: f column: 09  
Seq primer: 5' lambda-triPLEX2 Sequencing Primer.  
Location/Qualifiers  
1. 480

## FEATURES

source

/organism="Homo sapiens"  
/mol\_type="mrna"  
/db\_xref="taxon:9606"  
/clone="C149f09"  
/sex="unknown"  
/tissue\_type="blood"  
/cell\_type="Erythroid Precursor Cells"  
/cell\_line="Primary Culture of Peripheral Blood Mononuclear Cells"  
/dev\_stage="Precursor erythroblasts; GPA+"  
/lab\_host="DH5alpha"  
/clone\_lib="Hembase; Erythroid Precursor Cells (LCB:c1 library)"

/note="Organ: blood; Vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos (N-terminal )-biotin-GTC-CAC-CCG-AGG-CYT-G-(C-terminal) and (N-terminal)-biotin-C/T/C/T-GAA-GTT-CAT-AGG-A-(C-terminal). Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC;  
Http://www.nisc.nih.gov/)."

BASE COUNT 82 a 134 c 145 g 119 t  
ORIGIN

Query Match 31.8%; Score 399.2; DB 13; Length 480;  
Best Local Similarity 98.1%; Pred. No. 7.8e-94;  
Matches 404; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ATGAGCTTAAGTACCGGGGCTGTCGGCGCTGCTGCCCTCGGGCCCTTAACATG 60  
DB 69 ATGAGCTTAAGTACCGGGGCTGTCGGCGCTGCTGCCCTCGGGCCCTTAACATG 128  
QY 61 GAAGCAGCT 120  
DB 129 GAAGCAGCT 188  
QY 121 CAAAGGGGCTCGGCAATCTATCAAGTTGGCCAAATCTGACCGTGATGGCGCCATT 180  
DB 189 CAAAGGGGCTCGGCAATCTATCAAGTTGGCCAAATCTGACCGTGATGGCGCCATT 248  
QY 181 GGGTTGGGCTTCTCTCACTCGAGTTCCGAGACACAGCTGGAGCAGTGGGCTTCAAC 240  
DB 249 GGGTTGGGCTTCTCTCACTCGAGTTCCGAGACACAGCTGGAGCAGTGGGCTTCAAC 308  
QY 241 CTCCTCATCTGCGGCTTGGTGTGAGTGGGCAATCTCTGGACGGCTTCCAGACCA 300  
DB 309 CTCCTCATCTGCGGCTTGGTGTGAGTGGGCAATCTCTGGACGGCTTCCAGACCA 368  
QY 301 TTCCTTCTGGGAAGGTGTCATCACACTGTTCAGTATTGCGGCTGGCCACCATGAGTGCT 360  
DB 369 TTCCTTCTGGGAAGGTGTCATCACACTGTTCAGTATTGCGGCTGGCCACCATGAGTGCT 428  
QY 361 TTGTGGGTGCTGATCTCATGATGCTCTGTGGGGAAGTCAACTTTGGCCC 412  
DB 429 ATGTGGGTGCTGATCTCATGATGCTCTGTGGGGAAGTCAACTTTGGCCC 480

Search completed: September 26, 2003, 17:37:23  
Job time : 4808 secs